

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

(19) World Intellectual Property Organization
International Bureau



B3

(43) International Publication Date
23 October 2003 (23.10.2003)

PCT

(10) International Publication Number
WO 03/086308 A2

- (51) International Patent Classification⁷: **A61K**
- (21) International Application Number: PCT/US03/11231
- (22) International Filing Date: 9 April 2003 (09.04.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
10/121,988 11 April 2002 (11.04.2002) US
10/200,562 19 July 2002 (19.07.2002) US
10/237,551 6 September 2002 (06.09.2002) US
- (71) Applicant (*for all designated States except US*): **CORIXA CORPORATION** [US/US]; 1124 Columbia Street, Suite 200, Seattle, WA 98104 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): **DAY, Craig, H.** [US/US]; 19020 Wallingford Avenue N., Shoreline, WA 98133-4124 (US). **HOSKEN, Nancy, A.** [US/US]; 2314 North 56th Street, Seattle, WA 98103 (US). **PARSONS, Joseph, M.** [US/US]; 714 Bellevue Avenue E., Apt. 501, Seattle, WA 98104 (US).
- (74) Agents: **CHRISTIANSEN, William, T. et al.**; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— without international search report and to be republished upon receipt of that report
— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF HERPES SIMPLEX VIRUS INFECTION

(57) Abstract: Compounds and methods for the diagnosis and treatment of HSV infection are provided. The compounds comprise polypeptides that contain at least one antigenic portion of an HSV polypeptide and DNA sequences encoding such polypeptides. Pharmaceutical compositions and vaccines comprising such polypeptides or DNA sequences are also provided, together with antibodies directed against such polypeptides. Diagnostic kits are also provided comprising such polypeptides and/or DNA sequences and a suitable detection reagent for the detection of HSV infection in patients and in biological samples.

WO 03/086308 A2

COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF HERPES SIMPLEX VIRUS INFECTION

BACKGROUND OF THE INVENTION

Field of the Invention

5 The present invention relates generally to the detection and treatment of HSV infection. In particular, the invention relates to polypeptides comprising HSV antigens, DNA encoding HSV antigens, and the use of such compositions for the diagnosis and treatment of HSV infection.

Description of the Related Art

10 The herpes viruses include the herpes simplex viruses (HSV), comprising two closely related variants designated types 1 (HSV-1) and 2 (HSV-2). HSV is a prevalent cause of genital infection in humans, with an estimated annual incidence of 600,000 new cases and with 10 to 20 million individuals experiencing symptomatic chronic recurrent disease. The asymptomatic subclinical infection rate
15 may be even higher. For example, using a type-specific serological assay, 35% of an unselected population of women attending a health maintenance organization clinic in Atlanta had antibodies to HSV type 2 (HSV-2). Although continuous administration of antiviral drugs such as acyclovir ameliorates the severity of acute HSV disease and reduces the frequency and duration of recurrent episodes, such chemotherapeutic
20 intervention does not abort the establishment of latency nor does it alter the status of the latent virus. As a consequence, the recurrent disease pattern is rapidly reestablished upon cessation of drug treatment.

 The genome of at least one strain of herpes simplex virus (HSV) has been characterized. It is approximately 150 kb and encodes about 85 known genes,
25 each of which encodes a protein in the range of 50-1000 amino acids in length. Unknown, however, are the immunogenic portions, particularly immunogenic epitopes, that are capable of eliciting an effective T cell immune response to viral infection.

Thus, it is a matter of great medical and scientific need to identify immunogenic portions, preferably epitopes, of HSV polypeptides that are capable of eliciting an effective immune response to HSV infection. Such information will lead to safer and more effective prophylactic pharmaceutical compositions, *e.g.*, vaccine
5 compositions, to substantially prevent HSV infections, and, where infection has already occurred, therapeutic compositions to combat the disease. The present invention fulfills these and other needs.

BRIEF SUMMARY OF THE INVENTION

The present invention provides compositions and methods for the
10 diagnosis and therapy of HSV infection. In one aspect, the present invention provides polypeptides comprising an immunogenic portion of a HSV antigen, or a variant or biological functional equivalent of such an antigen. Certain preferred portions and other variants are immunogenic, such that the ability of the portion or variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments,
15 the polypeptide comprises an amino acid sequence encoded by a polynucleotide sequence selected from the group consisting of (a) a sequence of any one of SEQ ID NO: 1, 4, 8-9, 13, 16, 19 24, 35-38, 48-49, 52-53, 65-73, 76-89, 98-117, 118-119, 141, 144-152, 179-180 182-183, 184-194 206-210, 213-214, 217-226, 240, 242, 244-247, and 251-252; (b) a complement of said sequence; and (c) sequences that hybridize to a
20 sequence of (a) or (b) under moderately stringent conditions. In specific embodiments, the polypeptides of the present invention comprise at least a portion, preferably at least an immunogenic portion, of a HSV protein that comprises some or all of an amino acid sequence recited in any one of SEQ ID NO: 2, 3, 5, 6, 7, 10-12, 14-15, 17-18, 20-23, 25-33, 39-47, 50-51, 54-64, 74-75, 90-97, 120-121, 122-140, 142-143, 153-178, 181,
25 195-205 211-212, 215-216, 227-239, 241, 243, 248-250, 253-254, 255-267 including variants and biological functional equivalents thereof.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 contiguous amino acid residues of a HSV protein), expression vectors comprising

such polynucleotides and host cells transformed or transfected with such expression vectors.

In a related aspect, polynucleotide sequences encoding the above polypeptides, recombinant expression vectors comprising one or more of these
5 polynucleotide sequences and host cells transformed or transfected with such expression vectors are also provided.

In another aspect, the present invention provides fusion proteins comprising one or more HSV polypeptides, for example in combination with a physiologically acceptable carrier or immunostimulant for use as pharmaceutical
10 compositions and vaccines thereof.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody, either polyclonal and monoclonal, or antigen-binding fragment thereof that specifically binds to a HSV protein; and (b) a physiologically acceptable carrier.

15 Within other aspects, the present invention provides pharmaceutical compositions that comprise one or more HSV polypeptides or portions thereof disclosed herein, or a polynucleotide molecule encoding such a polypeptide, and a physiologically acceptable carrier. The invention also provides vaccines for prophylactic and therapeutic purposes comprising one or more of the disclosed polypeptides and an
20 immunostimulant, as defined herein, as well as vaccines comprising one or more polynucleotide sequences encoding such polypeptides and an immunostimulant.

In yet another aspect, methods are provided for inducing protective immunity in a patient, comprising administering to a patient an effective amount of one or more of the above pharmaceutical compositions or vaccines. Any of the polypeptides
25 identified for use in the treatment of patients can be used in conjunction with pharmaceutical agents used to treat herpes infections, such as, but not limited to, Zovirax®(Acyclovir), Valtrex® (Valacyclovir), and Famvir® (Famcyclovir).

In yet a further aspect, there are provided methods for treating, substantially preventing or otherwise ameliorating the effects of an HSV infection in a
30 patient, the methods comprising obtaining peripheral blood mononuclear cells (PBMC) from the patient, incubating the PBMC with a polypeptide of the present invention (or a

polynucleotide that encodes such a polypeptide) to provide incubated T cells and administering the incubated T cells to the patient. The present invention additionally provides methods for the treatment of HSV infection that comprise incubating antigen presenting cells with a polypeptide of the present invention (or a polynucleotide that
5 encodes such a polypeptide) to provide incubated antigen presenting cells and administering the incubated antigen presenting cells to the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient. In certain embodiments, the antigen presenting cells are selected from the group consisting of dendritic cells, macrophages, monocytes, B-cells, and fibroblasts. Compositions for the
10 treatment of HSV infection comprising T cells or antigen presenting cells that have been incubated with a polypeptide or polynucleotide of the present invention are also provided. Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

15 The present invention further provides, within other aspects, methods for removing HSV-infected cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a HSV protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

20 Within related aspects, methods are provided for inhibiting the development of HSV infection in a patient, comprising administering to a patient a biological sample treated as described above. In further aspects of the subject invention, methods and diagnostic kits are provided for detecting HSV infection in a patient. In one embodiment, the method comprises: (a) contacting a biological sample
25 with at least one of the polypeptides or fusion proteins disclosed herein; and (b) detecting in the sample the presence of binding agents that bind to the polypeptide or fusion protein, thereby detecting HSV infection in the biological sample. Suitable biological samples include whole blood, sputum, serum, plasma, saliva, cerebrospinal fluid and urine. In one embodiment, the diagnostic kits comprise one or more of the
30 polypeptides or fusion proteins disclosed herein in combination with a detection

reagent. In yet another embodiment, the diagnostic kits comprise either a monoclonal antibody or a polyclonal antibody that binds with a polypeptide of the present invention.

The present invention also provides methods for detecting HSV infection comprising: (a) obtaining a biological sample from a patient; (b) contacting the sample
5 with at least two oligonucleotide primers in a polymerase chain reaction, at least one of the oligonucleotide primers being specific for a polynucleotide sequence disclosed herein; and (c) detecting in the sample a polynucleotide sequence that amplifies in the presence of the oligonucleotide primers. In one embodiment, the oligonucleotide primer comprises at about 10 contiguous nucleotides of a polynucleotide sequence
10 peptide disclosed herein, or of a sequence that hybridizes thereto.

In a further aspect, the present invention provides a method for detecting HSV infection in a patient comprising: (a) obtaining a biological sample from the patient; (b) contacting the sample with an oligonucleotide probe specific for a polynucleotide sequence disclosed herein; and (c) detecting in the sample a
15 polynucleotide sequence that hybridizes to the oligonucleotide probe. In one embodiment, the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a polynucleotide sequence disclosed herein, or a sequence that hybridizes thereto.

These and other aspects of the present invention will become apparent
20 upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE SEVERAL SEQUENCE IDENTIFIERS

SEQ ID NO: 1 sets forth a polynucleotide sequence of an isolated clone
25 designated HSV2I_UL39fragH12A12;

SEQ ID NO: 2 sets forth an amino acid sequence, designated H12A12orf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 1;

SEQ ID NO: 3 sets forth the amino acid sequence of the full length
30 HSV-2 UL39 protein;

SEQ ID NO: 4 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_US8AfragD6.B_B11_T7Trc.seq;

5 SEQ ID NO: 5 sets forth an amino acid sequence, designated D6Borf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 4;

SEQ ID NO: 6 sets forth an amino acid sequence, designated D6Borf2.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 4;

10 SEQ ID NO: 7 sets forth the amino acid sequence of the full length HSV-2 US8A protein;

SEQ ID NO: 8 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_US4fragF10B3_T7Trc.seq;

SEQ ID NO: 9 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_US3fragF10B3_T7P.seq;

15 SEQ ID NO: 10 sets forth an amino acid sequence, designated F10B3orf2.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 8;

20 SEQ ID NO: 11 sets forth an amino acid sequence, designated 8F10B3orf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 9;

SEQ ID NO: 12 sets forth the amino acid sequence of the full length HSV-2 US3 protein;

SEQ ID NO: 13 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_UL46fragF11F5_T7Trc.seq

25 SEQ ID NO: 14 sets forth an amino acid sequence, designated F11F5orf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 13;

SEQ ID NO: 15 sets forth the amino acid sequence of the full length HSV-2 UL46 protein;

30 SEQ ID NO: 16 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_UL27fragH2C7_T7Trc.seq

SEQ ID NO: 17 sets forth an amino acid sequence, designated H2C7orf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 16;

5 SEQ ID NO: 18 sets forth the amino acid sequence of the full length HSV-2 UL27 protein;

SEQ ID NO: 19 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_UL18fragF10A1_rc.seq;

10 SEQ ID NO: 20 sets forth an amino acid sequence, designated F10A1orf3.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 19;

SEQ ID NO: 21 sets forth an amino acid sequence, designated F10A1orf2.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 19;

15 SEQ ID NO: 22 sets forth an amino acid sequence, designated F10A1orf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 19;

SEQ ID NO: 23 sets forth the amino acid sequence of the full length HSV-2 UL18 protein;

20 SEQ ID NO: 24 sets forth a polynucleotide sequence of an isolated clone designated HSV2II_UL15fragF10A12_rc.seq;

SEQ ID NO: 25 sets forth an amino acid sequence, designated F10A12orf1.pro, of an open reading frame encoded within the polynucleotide of SEQ ID NO: 24;

25 SEQ ID NO: 26 sets forth the amino acid sequence of the full length HSV-2 UL15 protein;

SEQ ID NO:27 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL46 gene;

SEQ ID NO:28 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL46 gene;

30 SEQ ID NO:29 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL46 gene;

SEQ ID NO:30 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL46 gene;

SEQ ID NO:31 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL46 gene;

5 SEQ ID NO:32 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL18 gene;

SEQ ID NO:33 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL18 gene;

10 SEQ ID NO:34 sets forth a nucleotide sequence of an isolated clone designated RL2_E9A4_5_consensus.seq;

SEQ ID NO:35 sets forth the nucleotide sequence of the full length HSV-2 RL2 gene;

SEQ ID NO:36 sets forth the nucleotide sequence of an isolated clone designated UL23_22_C12A12_consensus.seq;

15 SEQ ID NO:37 sets forth the nucleotide sequence of the full length HSV-2 UL23 protein;

SEQ ID NO:38 sets forth the nucleotide sequence of the full length HSV-2 UL22 protein;

20 SEQ ID NO:39 sets forth an amino acid sequence, designated HSV2_UL23, of an open reading frame encoded by the polynucleotide of SEQ ID NO: 37;

SEQ ID NO:40 sets forth an amino acid sequence designated HSV2_UL23 of an open reading frame encoded within the polynucleotides of SEQ ID NO:36;

25 SEQ ID NO:41 sets forth an amino acid sequence designated HSV2_UL22 of an open reading frame encoded within the polynucleotides of SEQ ID NO:36;

SEQ ID NO:42 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL23 gene;

30 SEQ ID NO:43 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL23 gene;

SEQ ID NO:44 sets forth the amino acid sequence of a 15-mer polypeptide derived from an immunogenic portion of the HSVII UL23 gene;

SEQ ID NO:45 sets forth an amino acid sequence, designated HSV2_UL22, of an open reading frame encoded by the polynucleotide of SEQ ID
5 NO:38;

SEQ ID NO:46 sets forth an amino acid sequence, designated RL2_E9A4_5_consensus.seq, of an open reading frame encoded by the polynucleotide of SEQ ID NO:34;

SEQ ID NO:47 sets forth an amino acid sequence, designated
10 HSV2_RL2, of an open reading frame encoded by the polynucleotide of SEQ ID NO:35;

SEQ ID NO:48 sets forth a nucleotide sequence of an isolated clone designated G10_UL37consensus.seq;

SEQ ID NO:49 sets forth the nucleotide sequence of the full length
15 HSV-2 UL37 gene;

SEQ ID NO:50 sets forth an amino acid sequence, designated HSV2_UL37, of an open reading frame encoded by the polynucleotide of SEQ ID NO:48; and

SEQ ID NO:51 sets forth an amino acid sequence, designated
20 HSV2_UL37, of an open reading frame encoded by the polynucleotide of SEQ ID NO:49;

SEQ ID NO:52 sets forth the DNA sequence derived from the insert of clone UL46fragF11F5;

SEQ ID NO:53 sets forth the DNA sequence derived from the insert of
25 clone G10;

SEQ ID NO:54 sets forth the amino acid sequence derived from the insert of clone UL46fragF11F5;

SEQ ID NO:55 sets forth the amino acid sequence derived from the insert of clone G10;

30 SEQ ID NO:56 is amino acid sequence of peptide #23 (amino acids 688-702) of the HSV-2 gene UL15;

SEQ ID NO:57 is amino acid sequence of peptide #30 (amino acids 716-730) of the HSV-2 gene UL15;

SEQ ID NO:58 is amino acid sequence of peptide #7 (amino acids 265-279) of the HSV-2 gene UL23;

5 SEQ ID NO:59 is amino acid sequence of peptide #2 (amino acids 621-635) of the HSV-2 gene UL46;

SEQ ID NO:60 is amino acid sequence of peptide #8 (amino acids 645-659) of the HSV-2 gene UL46;

10 SEQ ID NO:61 is amino acid sequence of peptide #9 (amino acids 649-663) of the HSV-2 gene UL46;

SEQ ID NO:62 is amino acid sequence of peptide #11 (amino acids 657-671) of the HSV-2 gene UL46;

SEQ ID NO:63 is amino acid sequence of peptide #33 (amino acids 262-276) of the HSV-2 gene US3;

15 SEQ ID NO:64 is amino acid sequence of peptide #5 (amino acids 99-113) of the HSV-2 gene US8A.

SEQ ID NO:65 sets forth the polynucleotide sequence of the full length HSV-2 UL39 protein.

20 SEQ ID NO:66 sets forth the partial polynucleotide sequence of UL39 derived from the HSV2-III library, pools 1F4, 1G2, and 3G11 which were recognized by clone 39.

SEQ ID NO:67 sets forth the partial polynucleotide sequence of UL39 derived from the HSV2-III library, pool 2C4 which was recognized by clone 39.

25 SEQ ID NO:68 sets forth the 5' end of the partial polynucleotide sequence of ICP0 derived from the HSV2-III library, pools 3H6, 3F12, and 4B2 which were recognized by clone 47.

SEQ ID NO:69 sets forth the 3' end of the partial polynucleotide sequence of ICP0 derived from the HSV2-III library, pools 3H6, 3F12, and 4B2 which were recognized by clone 47.

SEQ ID NO:70 sets forth the 5' end of the partial polynucleotide sequence of ICP0 derived from the HSV2-III library, pool 3A1 which was recognized by clone 47.

5 SEQ ID NO:71 sets forth the 3' end of the partial polynucleotide sequence of ICP0 derived from the HSV2-III library, pool 3A1 which was recognized by clone 47.

SEQ ID NO:72 sets forth the 5' end of the partial polynucleotide sequence of ICP0 derived from the HSV2-III library, pool 2B2 which was recognized by clone 47.

10 SEQ ID NO:73 sets forth the 3' end of the partial polynucleotide sequence of ICP0 derived from the HSV2-III library, pool 2B2 which was recognized by clone 47.

15 SEQ ID NO:74 sets forth the partial amino acid sequence of UL39 derived from the HSV2-III library, pools 1F4, 1G2, and 3G11 which were recognized by clone 39.

SEQ ID NO:75 sets forth the partial amino acid sequence of UL39 derived from the HSV2-III library, pool 2C4 which was recognized by clone 39.

SEQ ID NO:76 sets forth a full length DNA sequence for the HSV-2 gene UL19.

20 SEQ ID NO:77 sets forth a DNA sequence for the vaccinia virus shuttle plasmid, pSC11.

SEQ ID NO:78 sets forth a full length DNA sequence for the HSV-2 gene, UL47.

25 SEQ ID NO:79 sets forth a full length DNA sequence for the HSV-2 gene, UL50.

SEQ ID NO:80 sets forth a DNA sequence for the human Ubiquitin gene.

SEQ ID NO:81 sets forth a full length DNA sequence for the HSV-2 gene, UL49.

30 SEQ ID NO:82 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL50.

SEQ ID NO:83 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL49.

SEQ ID NO:84 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL19.

5 SEQ ID NO:85 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL21.

SEQ ID NO:86 sets forth a DNA sequence corresponding to the coding region of the HSV-2 UL47 gene with the Trx2 fusion sequence.

10 SEQ ID NO:87 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL47.

SEQ ID NO:88 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL47 C fragment.

SEQ ID NO:89 sets forth a DNA sequence corresponding to the coding region of the HSV gene, UL39.

15 SEQ ID NO:90 sets forth an amino acid sequence corresponding to the UL39 protein with a His tag.

SEQ ID NO:91 sets forth an amino acid sequence corresponding to the UL21 protein with a His tag.

20 SEQ ID NO:92 sets forth an amino acid sequence corresponding to the UL47 protein fused with the Trx and 2 histadine tags.

SEQ ID NO:93 sets forth an amino acid sequence corresponding to the UL47 C fragment with a His tag.

SEQ ID NO:94 sets forth an amino acid sequence corresponding to the UL47 protein with a His tag.

25 SEQ ID NO:95 sets forth an amino acid sequence corresponding to the UL19 protein with a His tag.

SEQ ID NO:96 sets forth an amino acid sequence corresponding to the UL50 protein with a His tag.

30 SEQ ID NO:97 sets forth an amino acid sequence corresponding to the UL49 protein with a His tag.

SEQ ID NO:98 sets forth the primer sequence for the sense primer PDM-602, used in the amplification of UL21.

SEQ ID NO:99 sets forth the primer sequence for the reverse primer PDM-603, used in the amplification of UL21.

5 SEQ ID NO:100 sets forth the primer sequence for the sense primer PDM-466, used in the amplification of UL39.

SEQ ID NO:101 sets forth the primer sequence for the reverse primer PDM-467, used in the amplification of UL39.

10 SEQ ID NO:102 sets forth the primer sequence for the sense primer PDM-714, used in the amplification of UL49.

SEQ ID NO:103 sets forth the primer sequence for the reverse primer PDM-715, used in the amplification of UL49.

SEQ ID NO:104 sets forth the primer sequence for the sense primer PDM-458, used in the amplification of UL50.

15 SEQ ID NO:105 sets forth the primer sequence for the reverse primer PDM-459, used in the amplification of UL50.

SEQ ID NO:106 sets forth the primer sequence for the sense primer PDM-453, used in the amplification of UL19.

20 SEQ ID NO:107 sets forth the primer sequence for the reverse primer PDM-457, used in the amplification of UL19.

SEQ ID NO:108 sets forth the primer sequence for the sense primer PDM-631, used in the amplification of UL47.

SEQ ID NO:109 sets forth the primer sequence for the reverse primer PDM-632, used in the amplification of UL47.

25 SEQ ID NO:110 sets forth the primer sequence for the sense primer PDM-631, used in the amplification of UL47 A.

SEQ ID NO:111 sets forth the primer sequence for the reverse primer PDM-645, used in the amplification of UL47 A.

30 SEQ ID NO:112 sets forth the primer sequence for the sense primer PDM-646, used in the amplification of UL47 B.

SEQ ID NO:113 sets forth the primer sequence for the reverse primer PDM-632, used in the amplification of UL47 B.

SEQ ID NO:114 sets forth the primer sequence for the sense primer PDM-631, used in the amplification of UL47 C.

5 SEQ ID NO:115 sets forth the primer sequence for the reverse primer PDM-739, used in the amplification of UL47 C.

SEQ ID NO:116 sets forth the primer sequence for the sense primer PDM-740, used in the amplification of UL47 D.

10 SEQ ID NO:117 sets forth the primer sequence for the reverse primer PDM-632, used in the amplification of UL47 D.

SEQ ID NO:118 sets forth a novel DNA sequence for the HSV-2 gene, US8.

SEQ ID NO:119 sets forth the published DNA sequence for the HSV-2 gene, US8, derived from the HG52 strain of HSV-2.

15 SEQ ID NO:120 sets forth an amino acid sequence encoded by SEQ ID NO:118.

SEQ ID NO:121 sets forth an amino acid sequence encoded by SEQ ID NO:119.

20 SEQ ID NO:122 sets forth the sequence of peptide 85 (p85), a CD8+ peptide derived from the HSV-2 gene, UL47.

SEQ ID NO:123 sets forth the sequence of peptide 89 (p89), a CD8+ peptide derived from the HSV-2 gene, UL47.

SEQ ID NO:124 sets forth the sequence of peptide 98/99 (p98/99), a CD8+ peptide derived from the HSV-2 gene, UL47.

25 SEQ ID NO:125 sets forth the sequence of peptide 105 (p105), a CD8+ peptide derived from the HSV-2 gene, UL47.

SEQ ID NO:126 sets forth the sequence of peptide 112 (p112), a CD8+ peptide derived from the HSV-2 gene, UL47.

30 SEQ ID NO:127 sets forth the sequence of peptide #23 (amino acids 688-702) from the HSV-2 protein UL15.

SEQ ID NO:128 sets forth the sequence of peptide #30 (amino acids 716-730) from the HSV-2 protein UL15.

SEQ ID NO:129 sets forth the sequence of peptide #7 (amino acids 265-272) from the HSV-2 protein UL23.

5 SEQ ID NO:130 sets forth the sequence of peptide #2 (amino acids 621-635) from the HSV-2 protein UL46.

SEQ ID NO:131 sets forth the sequence of peptide #8 (amino acids 645-659) from the HSV-2 protein UL46.

10 SEQ ID NO:132 sets forth the sequence of peptide #9 (amino acids 649-663) from the HSV-2 protein UL46.

SEQ ID NO:133 sets forth the sequence of peptide #11 (amino acids 657-671) from the HSV-2 protein UL46.

SEQ ID NO:134 sets forth the sequence of peptide #86 (amino acids 341-355) from the HSV-2 protein UL47.

15 SEQ ID NO:135 sets forth the sequence of peptide #6 (amino acids 21-35) from the HSV-2 protein UL49.

SEQ ID NO:136 sets forth the sequence of peptide #12 (amino acids 45-59) from the HSV-2 protein UL49.

20 SEQ ID NO:137 sets forth the sequence of peptide #13 (amino acids 49-63) from the HSV-2 protein UL49.

SEQ ID NO:138 sets forth the sequence of peptide #49 (amino acids 193-208) from the HSV-2 protein UL49.

SEQ ID NO:139 sets forth the sequence of peptide #33 (amino acids 262-276) from the HSV-2 protein US3.

25 SEQ ID NO:140 sets forth the sequence of peptide #5 (amino acids 99-113) from the HSV-2 protein US8A.

SEQ ID NO:141 sets forth a full length insert DNA sequence corresponding to the clone F10B3.

30 SEQ ID NO:142 sets forth a full length insert amino acid sequence corresponding to the clone F10B3.

SEQ ID NO:143 sets forth an amino acid sequence for the HSV-2 protein, US4.

SEQ ID NO:144 sets forth a DNA sequence for the HSV-2 protein, UL21.

5 SEQ ID NO:145 sets forth a DNA sequence for the HSV-2 protein, UL50.

SEQ ID NO:146 sets forth a DNA sequence for the HSV-2 protein, US3.

SEQ ID NO:147 sets forth a DNA sequence for the HSV-2 protein, UL54.

10 SEQ ID NO:148 sets forth a DNA sequence for the HSV-2 protein, US8.
SEQ ID NO:149 sets forth a DNA sequence for the HSV-2 protein, UL19.

SEQ ID NO:150 sets forth a DNA sequence for the HSV-2 protein, UL46.

15 SEQ ID NO:151 sets forth a DNA sequence for the HSV-2 protein, UL18.

SEQ ID NO:152 sets forth a DNA sequence for the HSV-2 protein, RL2.

SEQ ID NO:153 sets forth an amino sequence for the HSV-2 protein, UL50.

20 SEQ ID NO:154 sets forth an amino acid sequence for the HSV-2 protein, UL21.

SEQ ID NO:155 sets forth an amino acid sequence for the HSV-2 protein, US3.

25 SEQ ID NO:156 sets forth an amino acid sequence for the HSV-2 protein, UL54.

SEQ ID NO:157 sets forth an amino acid sequence for the HSV-2 protein, US8.

SEQ ID NO:158 sets forth an amino acid sequence for the HSV-2 protein, UL19.

30 SEQ ID NO:159 sets forth an amino acid sequence for the HSV-2 protein, UL46.

SEQ ID NO:160 sets forth an amino acid sequence for the HSV-2 protein, UL18.

SEQ ID NO:161 sets forth an amino acid sequence for the HSV-2 protein, RL2.

5 SEQ ID NO:162 sets forth the sequence of peptide #43 (amino acids 211-225) from the HSV-2 protein RL2.

SEQ ID NO:163 sets forth the sequence of peptide #41 (amino acids 201-215) from the HSV-2 protein UL46.

10 SEQ ID NO:164 sets forth the sequence of peptide #50 (amino acids 246-260) from the HSV-2 protein UL46.

SEQ ID NO:165 sets forth the sequence of peptide #51 (amino acids 251-265) from the HSV-2 protein UL46.

SEQ ID NO:166 sets forth the sequence of peptide #60 (amino acids 296-310) from the HSV-2 protein UL46.

15 SEQ ID NO:167 sets forth the sequence of peptide #74 (amino acids 366-380) from the HSV-2 protein US8.

SEQ ID NO:168 sets forth the sequence of peptide #102 (amino acids 506-520) from the HSV-2 protein UL19.

20 SEQ ID NO:169 sets forth the sequence of peptide #103 (amino acids 511-525) from the HSV-2 protein UL19.

SEQ ID NO:170 sets forth the sequence of peptide #74 (amino acids 366-380) from the HSV-2 protein UL19.

SEQ ID NO:171 sets forth the sequence of peptide #75 (amino acids 371-385) from the HSV-2 protein UL19.

25 SEQ ID NO:172 sets forth the sequence of peptide #17 (amino acids 65-79) from the HSV-2 protein UL18.

SEQ ID NO:173 sets forth the sequence of peptide #18 (amino acids 69-83) from the HSV-2 protein UL18.

30 SEQ ID NO:174 sets forth the sequence of peptide #16 (amino acids 76-90) from the HSV-2 protein UL50.

SEQ ID NO:175 sets forth the sequence of peptide #23 (amino acids 111-125) from the HSV-2 protein UL50.

SEQ ID NO:176 sets forth the sequence of peptide #49 (amino acids 241-255) from the HSV-2 protein UL50.

5 SEQ ID NO:177 sets forth the sequence of a 9-mer peptide for ICP0 (amino acids 215-223).

SEQ ID NO:178 sets forth the sequence of a 10-mer peptide for UL46 (amino acids 251-260).

10 SEQ ID NO:179 sets forth a DNA sequence of US4 derived from the HG52 strain of HSV-2.

SEQ ID NO:180 sets forth a DNA sequence for the UL47 F coding region.

SEQ ID NO:181 sets forth an amino acid sequence for the UL47 F coding region.

15 SEQ ID NO:182 sets forth the sequence for primer CBH-002 used in the amplification of UL47 F.

SEQ ID NO:183 sets forth the sequence for primer PDM-632 used in the amplification of UL47 F.

20 SEQ ID NO:184 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL18.

SEQ ID NO:185 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, LAT-ORF-1.

SEQ ID NO:186 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL48.

25 SEQ ID NO:187 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL41.

SEQ ID NO:188 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL39.

30 SEQ ID NO:189 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL37.

SEQ ID NO:190 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL36.

SEQ ID NO:191 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL29.

5 SEQ ID NO:192 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL25.

SEQ ID NO:193 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, ICP4.

10 SEQ ID NO:194 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, ICP22.

SEQ ID NO:195 sets forth a full length amino acid sequence corresponding to the HSV-2 open reading frame, UL18.

SEQ ID NO:196 sets forth a full length amino acid sequence corresponding to the HSV-2 open reading frame, ICP22.

15 SEQ ID NO:197 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, ICP4.

SEQ ID NO:198 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, LAT-ORF-1.

20 SEQ ID NO:199 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL25.

SEQ ID NO:200 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL29.

SEQ ID NO:201 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL36.

25 SEQ ID NO:202 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL37.

SEQ ID NO:203 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL39.

30 SEQ ID NO:204 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL41.

SEQ ID NO:205 sets forth a full length DNA sequence corresponding to the HSV-2 open reading frame, UL48.

SEQ ID NO:206 sets forth the DNA sequence from the E. coli expression cloning library inserts 1/F3 and 1/A7.

5 SEQ ID NO:207 sets forth the DNA sequence from the E. coli expression cloning library insert 1/H6.

SEQ ID NO:208 sets forth the DNA sequence from the E. coli expression cloning library insert 3/C1.

10 SEQ ID NO:209 sets forth the DNA sequence common to the inserts 1/F3, 1/A7, 1/H6, and 3/C1.

SEQ ID NO:210 sets forth a full length DNA sequence for the UL19 gene derived from HSV-2 strain HG52.

SEQ ID NO:211 sets forth the amino acid sequence encoded by SEQ ID NO:209.

15 SEQ ID NO:212 sets forth a full length amino acid sequence for the UL19 gene derived from the HSV-2 strain HG52.

SEQ ID NO:213 sets forth the full length DNA sequence for the US8 gene derived from the clinical isolate RW1874.

20 SEQ ID NO:214 sets forth the full length DNA sequence for the US8 gene derived from the clinical isolate HV5101.

SEQ ID NO:215 sets forth the full length amino acid sequence for the US8 gene derived from the clinical isolate RW1874.

SEQ ID NO:216 sets forth the full length amino acid sequence for the US8 gene derived from the clinical isolate HV5101.

25 SEQ ID NO:217 sets forth the DNA sequence corresponding to the insert from clone HH6 D6_B6.

SEQ ID NO:218 sets forth a DNA sequence for the UL21 gene derived from the HSV-2 strain HG52.

30 SEQ ID NO:219 sets forth a DNA sequence corresponding to the first insert of clone HH20 C12_E1.

SEQ ID NO:220 sets forth a DNA sequence corresponding to the second insert of clone HH20 C12_E1.

SEQ ID NO:221 sets forth a DNA sequence for the UL29 gene derived from the HSV-2 strain HG52.

5 SEQ ID NO:222 sets forth the DNA sequence corresponding to the insert from clone HH22 F7_A7.

SEQ ID NO:223 sets forth the DNA sequence corresponding to the insert from clone HH22 4/E8_C8.

10 SEQ ID NO:224 sets forth a DNA sequence for the UL46 gene derived from the HSV-2 strain HG52.

SEQ ID NO:225 sets forth a DNA sequence corresponding to the insert from the clone HH24 G6_H11.

SEQ ID NO:226 sets forth a DNA sequence for the UL47 gene derived from the HSV-2 strain HG52.

15 SEQ ID NO:227 sets forth a protein sequence encoded by the insert from clone HH6 (D6B6:SEQ ID NO:217).

SEQ ID NO:228 sets forth a full length amino acid sequence for UL21 derived from the HG52 strain of HSV-2.

20 SEQ ID NO:229 sets forth an amino acid sequence of the UL21 T cell epitope spanning amino acids 281 to 300.

SEQ ID NO:230 sets forth an amino acid sequence encoded by the insert 1C12_E1, from clone HH20.

SEQ ID NO:231 sets forth an amino acid sequence encoded by the insert 2E9_D11, from clone HH20.

25 SEQ ID NO:232 sets forth a full-length amino acid sequence for the HSV-2, strain HG52 protein, UL29.

SEQ ID NO:233 sets forth an amino acid sequence from insert F7_A1, clone HH22.

30 SEQ ID NO:234 sets forth an amino acid sequence from insert 4/E8_C8, clone HH22.

SEQ ID NO:235 sets forth a full-length amino acid sequence for the UL46 protein derived from the HG52 strain of HSV-2.

SEQ ID NO:236 sets forth an amino acid sequence of the reactive T cell epitope derived from UL46, spanning amino acids 621 to 649.

5 SEQ ID NO:237 sets forth an amino acid sequence encoded by the insert derived from clone HH24 G6_H11.

SEQ ID NO:238 sets forth a full-length amino acid sequence for the HSV-2 gene UL47.

10 SEQ ID NO:239 sets forth an amino acid sequence of the reactive T cell epitope derived from UL47 spanning amino acids 137-155.

SEQ ID NO:240 sets forth a DNA sequence corresponding to the clone insert TM13 and TM58 F5_G1.

SEQ ID NO:241 sets forth an amino acid sequence encoded by the clone insert TM13 and TM58 F5_G1.

15 SEQ ID NO:242 sets forth a full length DNA sequence corresponding to the HSV-2 gene, UL54.

SEQ ID NO:243 sets forth a full length amino acid sequence corresponding to UL54 (ICP27) derived from the HG52 strain of HSV-2.

20 SEQ ID NO:244 sets forth a DNA sequence corresponding to the insert TM39, H11_C3.

SEQ ID NO:245 sets forth a full length DNA sequence corresponding to UL21, derived from the HSV-2 strain, HG52.

SEQ ID NO:246 sets forth a full length DNA sequence corresponding to UL22, derived from the HSV-2 strain, HG52.

25 SEQ ID NO:247 sets forth a full length DNA sequence corresponding to UL36, derived from the HSV-2 strain, HG52.

SEQ ID NO:248 sets forth a full length amino acid sequence corresponding to UL21 derived from the HG52 strain of HSV-2.

30 SEQ ID NO:249 sets forth a full length amino acid sequence corresponding to UL22 derived from the HG52 strain of HSV-2.

SEQ ID NO:250 sets forth a full length amino acid sequence corresponding to UL36 derived from the HG52 strain of HSV-2.

SEQ ID NO:251 sets forth a DNA sequence corresponding to the insert of TM51, F7_A8.

5 SEQ ID NO:252 sets forth a full length DNA sequence corresponding to US4, derived from the HSV-2 strain, HG52.

SEQ ID NO:253 sets forth an amino acid sequence corresponding to clone TM51 F7_A8.

10 SEQ ID NO:254 sets forth a full length amino acid sequence corresponding to US4 derived from the HG52 strain of HSV-2.

SEQ ID NO:255 sets forth the sequence of peptide #43 amino acid residues 211-225 of ICP0.

SEQ ID NO:256 sets forth the sequence of peptide #44 amino acid residues 216-230 of ICP0.

15 SEQ ID NO:257 sets forth the minimum CD8 epitope of ICP0 peptides #43 and #44.

SEQ ID NO:258 sets forth the sequence of peptide #89 amino acid residues 440-454 of ICP0.

20 SEQ ID NO:259 sets forth the sequence of peptide #82 amino acid residues 405-419 of ICP0.

SEQ ID NO:260 sets forth the sequence of peptide #79 amino acid residues 390-404 of ICP0.

SEQ ID NO:261 sets forth the sequence of peptide #98 amino acid residues 389-403 of UL47.

25 SEQ ID NO:262 sets forth the sequence of peptide #99 amino acid residues 393-407 of UL47.

SEQ ID NO:263 sets forth the minimum CD8 epitope of UL47 peptides #98 and #99.

30 SEQ ID NO:264 sets forth the sequence of peptide #112 amino acid residues 445-459 of UL47.

SEQ ID NO:265 sets forth the putative CD8 epitope of UL47 peptide #112.

SEQ ID NO:266 sets forth the sequence of peptide #4 amino acid residues 13-27 of UL47.

5 SEQ ID NO:267 sets forth the sequence of peptide #1 amino acid residues 1-15 of UL47.

DETAILED DESCRIPTION OF THE INVENTION

U.S. patents, U.S. patent application publications, U.S. patent applications, foreign patents, foreign patent applications and non-patent publications
10 referred to in this specification and/or listed in the Application Data Sheet, are incorporated herein by reference, in their entirety.

As noted above, the present invention is generally directed to compositions and methods for making and using the compositions, particularly in the therapy and diagnosis of HSV infection. Certain illustrative compositions described
15 herein include HSV polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (e.g., T cells). Certain HSV proteins and immunogenic portions thereof comprise HSV polypeptides that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient infected with HSV.

20 Therefore, the present invention provides illustrative polynucleotide compositions, illustrative polypeptide compositions, immunogenic portions of said polynucleotide and polypeptide compositions, antibody compositions capable of binding such polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human HSV
25 infections.

POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment

that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller fragments of such segments, and also recombinant vectors, 5 including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally 10 isolated, or modified synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA 15 segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, 20 which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

25 Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes an HSV protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded 30 polypeptide is not diminished, relative to a native HSV protein. The effect on the

immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term "variants" also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A "comparison window" as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O., *A model of evolutionary change in proteins – Matrices for detecting distant relationships*, 1978. In Dayhoff, M.O. (ed.), *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington D.C., Vol. 5, Suppl. 3, pp. 345-358; Hein.J., *Unified Approach to Alignment and Phylogenies*, "Methods in Enzymology," Academic Press, Inc., San Diego, CA vol. 183, pp. 626-645, 1990; Higgins, D.G. and P.M. Sharp, *CABIOS* 5:151-53, 1989; Myers, E.W. and W. Muller, *CABIOS* 4:11-17, 1988; Robinson, E.D., *Comb. Theor* 11:105, 1971; Santou, N. and M. Nes, *Mol. Biol. Evol.* 4:406-25, 1987; Sneath, P.H.A. and R.R. Sokal, *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA, 1973; Wilbur, W.J. and D.J. Lipman, *Proc. Natl. Acad. Sci. USA* 80:726-30, 1983.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman, *Add. APL. Math* 2:482, 1981, by the identity alignment algorithm of Needleman and Wunsch, *J. Mol. Biol.* 48:443, 1970, by the search for similarity methods of Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics

Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., *Nucl. Acids Res.* 25:3389-3402, 1977; and Altschul et al., *J. Mol. Biol.* 215:403-10, 1990, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915, 1989) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the

total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000,

about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like, (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

25 PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence

disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

5 The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

10 Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene
15 product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region
20 may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

 The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules
25 having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where
30 desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in the sequences disclosed herein, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one
5 wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly
10 practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCRTM technology of U.S. Patent No. 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art
15 of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of
20 selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate
25 little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be
30 needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M

salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to
5 destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using
10 any of a variety of well established techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for HSV-associated expression (*i.e.*, expression that is at least two fold greater in infected versus normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA)
15 according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this
20 approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.*, an HSV cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is
25 screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by
30 nick-translation or end-labeling with ³²P) using well known techniques. A bacterial or

bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are
5 selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may
10 involve generating a series of deletion clones. The resulting overlapping sequences can then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques,
15 amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be
20 sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (see Triglia et al., *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation
25 and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known
30 region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO

96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom et al., *PCR Methods Applic. 1*:111-19, 1991) and walking PCR (Parker et al., *Nucl. Acids. Res. 19*:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (e.g., NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide

encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide
5 sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to
10 encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved
15 and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. et al. (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical
20 methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. et al. (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

25 A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation
30 procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof,

may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. et al. (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses

are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant

cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

5 An insect system may also be used to express a polypeptide of interest. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control
10 of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. et al. (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

15 In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used
20 to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

 Specific initiation signals may also be used to achieve more efficient
25 translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion
30 thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct

reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or apt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which

confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. et al (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells
5 to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to
10 quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. et al. (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a
15 marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

20 Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification
25 of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated
30 cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be

preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) and Maddox, D. E. et al. (1983; *J. Exp. Med.* 158:1211-1216).

5 A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions
10 thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used
15 include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained
20 intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of
25 interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity
30 purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego,

Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine
5 residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. et al. (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. et al. (1993; *DNA Cell Biol.* 12:441-453).

10 In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide
15 Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of
20 individual peptides, or biologically functional equivalent polypeptides, through specific mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA.
25 Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected
30 polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the

properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing

potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence
5 variants. Specific details regarding these methods and protocols are found in the teachings of Maloy et al., 1994; Segal, 1976; Prokop and Bajpai, 1991; Kubby, 1994; and Maniatis et al., 1982, each incorporated herein by reference, for that purpose.

As used herein, the term "oligonucleotide directed mutagenesis procedure" refers to template-dependent processes and vector-mediated propagation
10 which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term "oligonucleotide directed mutagenesis procedure" is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent
15 process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of
20 the amplified nucleic acid fragment. Examples of such methodologies are provided by U.S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification
25 methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture
30 along with a DNA polymerase (e.g., *Taq* polymerase). If the target sequence is present

in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCRTM amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCRTM, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α -thio]triphosphates in one strand of a restriction site (Walker et al., 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand

displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh et al., 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques

involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer,
5 followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

10 Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase
15 (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is
20 then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies
25 can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by
30 reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA

("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.*, new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

5 Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in its entirety), may also be used in the amplification of DNA sequences of the present invention.

10 BIOLOGICAL FUNCTIONAL EQUIVALENTS

 Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific
15 polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

 When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid
20 changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

 For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites
25 on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide

sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE I

Amino Acids		Codons							
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			
Glutamine	Gln	Q	CAA	CAG					
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
Threonine	Thr	T	ACA	ACC	ACG	ACU			
Valine	Val	V	GUA	GUC	GUG	GUU			
Tryptophan	Trp	W	UGG						
Tyrosine	Tyr	Y	UAC	UAU					

5

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative
 10 hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other

molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine 5 (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with 10 similar biological activity, *i.e.*, still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 15 No. 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U.S. Patent No. 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); 20 aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (−0.4); proline (−0.5 \pm 1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain 25 a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based 30 on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that

take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase
5 stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl-methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and
10 uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety of well known approaches, several of which are outlined
15 below for the purpose of illustration.

1. ADENOVIRUS

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences
20 sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an
25 adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity.

Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease such as acute respiratory disease in humans.

5 Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different
10 transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression and host cell shut-off (Renan, 1990).
15 The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-tripartite leader (TPL) sequence which makes them preferred mRNA's for
20 translation.

 In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process. Therefore, it is critical to isolate a single clone of virus
25 from an individual plaque and examine its genomic structure.

 Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham et al., 1977). Since the E3 region is
30 dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the

D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury et al., 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher et al. (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fiba-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not

believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is
5 a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most
10 convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors as described by Karlsson et al.
15 (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, e.g., 10^9 - 10^{11} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus
20 does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch et al., 1963; Top et al., 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero et al., 1991; Gomez-Foix et al., 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet et al., 1990; Rich et al., 1993). Studies in
30 administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld et al., 1991; Rosenfeld et al., 1992), muscle injection (Ragot et al., 1993),

peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle et al., 1993).

2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann et al., 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann et al., 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind et al., 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

5 A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux et al., 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection
10 of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux et al., 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzyczka, 1984) is a parovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies
15 are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of
20 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs (FIG. 2). There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped
25 hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins,

and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar et al., 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar et al., 1988; Horwich et al., 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich et al., 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive properties for liver-directed gene transfer. Chang et al. (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B

virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days
5 after transfection (Chang et al., 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for
10 transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be
15 positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be
20 stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

25 In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but
30 it may be applied to *in vivo* use as well. Dubensky et al. (1984) successfully injected

polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes.

- 5 It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity
10 allowing them to pierce cell membranes and enter cells without killing them (Klein et al., 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang et al., 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

- 15 Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang et al., 1990; Zelenin et al., 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e.*, *ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present
20 invention.

ANTISENSE OLIGONUCLEOTIDES

- The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the
25 route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense

nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U.S. Patent No. 5,739,119 and U.S. Patent No. 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski et al., 1988; Vasanthakumar and Ahmed, 1989; Peris et al., 1998; U.S. Patent No. 5,801,154; U.S. Patent No. 5,789,573; U.S. Patent No. 5,718,709 and U.S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, e.g., cancer (U.S. Patent No. 5,747,470; U.S. Patent No. 5,591,317 and U.S. Patent No. 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.*, in these illustrative examples

the rat and human sequences) and determination of secondary structure, T_m , binding energy, relative stability, and antisense compositions were selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

5 Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm
10 software (Altschul et al., 1997).

 The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear-localization sequence of SV40 T-antigen (Morris et al., 1997).
15 It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris et al., 1997).

20 RIBOZYMES

 Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity
25 (Kim and Cech, 1987; Gerlach et al., 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech et al., 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via

specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech et al., 1981). For example, U.S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon et al., 1991; Sarver et al., 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes *H-ras*, *c-fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of

target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a
5 ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf et al., 1992). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA
10 guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi et al. (1992). Examples of hairpin motifs are described by Hampel et al. (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel et al. (1990) and U.S. Patent No. 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an
15 example of the RNaseP motif is described by Guerrier-Takada et al. (1983); Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U.S. Patent No. 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it
20 has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic
25 cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or
30 RNA vectors that are delivered to specific cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells from eukaryotic promoters (*e.g.*, Scanlon et al., 1991; Kashani-Sabet et al., 1992; Dropulic et al., 1992; Weerasinghe et al., 1991; Ojwang et al., 1992; Chen et al., 1992; Sarver et al., 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No. WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa et al., 1992; Taira et al., 1991; and Ventura et al., 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger et al., 1989) to assess whether the ribozyme sequences fold into the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman et al. (1987) and in Scaringe et al. (1990) and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-fluoro, 2'-o-methyl, 2'-H (for a review see, *e.g.*, Usman and Cedergrén, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see, *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al.*, 1990; Pieken *et al.*, 1991; Usman and Cedergrén, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U.S. Patent No. 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other

routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 5 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for 10 eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the 15 prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber et al., 1993; Zhou et al., 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.*, Kashani-Saber et al., 1992; Ojwang et al., 1992; Chen et al., 1992; Yu et al., 1993; L'Huillier et al., 1992; Lisiewicz et al., 1993). Such transcription units can be 20 incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and 25 mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA 30 structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially)

of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple ribozymes targeted to different genes, ribozymes coupled with
5 known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard
10 methodology.

PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and
15 Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference.
20 As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

25 PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen et al., 1991; Hanvey et al., 1992; Hyrup and Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a
30 stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm et al.,

1994) or Fmoc (Thomson et al., 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used (Christensen et al., 1995).

PNA monomers or ready-made oligomers are commercially available
5 from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton et al., 1995). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will
10 depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed
15 by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton et al., 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that
20 contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and
25 utilized modifications of PNAs (Norton et al., 1995; Haaima et al., 1996; Stetsenko et al., 1996; Petersen et al., 1995; Ulmann et al., 1996; Koch et al., 1995; Orum et al., 1995; Footer et al., 1996; Griffith et al., 1995; Kremsky et al., 1996; Pardridge et al., 1995; Boffa et al., 1995; Landsdorp et al., 1996; Gambacorti-Passerini et al., 1996; Armitage et al., 1997; Seeger et al., 1997; Ruskowski et al., 1997). U.S. Patent No.
30 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics,

modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs
5 recognize complementary DNA and RNA by Watson-Crick pairing (Egholm et al., 1993), validating the initial modeling by Nielsen et al. (1991). PNAs lack 3' to 5' polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm et al., 1993).

Hybridization of DNA oligonucleotides to DNA and RNA is destabilized
10 by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen et al., 1991). The enhanced rate and affinity of hybridization are significant because they are responsible
15 for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang et al. have shown that support-bound PNAs can be
20 used to detect hybridization events (Wang et al., 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence specificity of PNA recognition. As with DNA hybridization, however, selective recognition can be achieved by balancing oligomer length and incubation
25 temperature. Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a single mismatch within a 16 bp PNA-DNA duplex can reduce the T_m by up to 15°C (Egholm et al., 1993). This high level of discrimination has allowed the development of several PNA-based strategies for the analysis of point mutations (Wang
30 et al., 1996; Carlsson et al., 1996; Thiede et al., 1996; Webb and Hurskainen, 1996; Perry-O'Keefe et al., 1996).

High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an essential RNA template, while the analogous DNA oligomers do not (Norton et al., 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini (Nielsen et al., 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur spontaneously at sequences within chromosomal DNA (Boffa et al., 1995; Boffa et al., 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa et al., 1995) and to inhibit transcription (Boffa et al., 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene expression. Studies and reviews concerning the use of PNAs as antisense and anti-gene agents include Nielsen et al. (1993b), Hanvey et al. (1992), and Good and Nielsen (1997). Koppelhus et al. (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen et al. (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen et al. using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen et al., 1991), antisense inhibition (Hanvey et al., 1992), mutational analysis (Orum et al., 1993), enhancers of transcription (Mollegaard et al., 1994), nucleic acid purification (Orum et al., 1995), isolation of transcriptionally active genes (Boffa et al., 1995), blocking of transcription factor binding (Vickers et al., 1995), genome cleavage

(Veselkov et al., 1996), biosensors (Wang et al., 1996), *in situ* hybridization (Thisted et al., 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

The present invention, in other aspects, provides polypeptide
5 compositions. Generally, a polypeptide of the invention will be an isolated polypeptide
(or an epitope, variant, or active fragment thereof) derived from HSV. Preferably, the
polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence
which hybridizes under moderate or highly stringent conditions to a polynucleotide
sequence disclosed herein. Alternatively, the polypeptide may be defined as a
10 polypeptide which comprises a contiguous amino acid sequence from an amino acid
sequence disclosed herein, or which polypeptide comprises an entire amino acid
sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to
comprise one or more polypeptides that are immunologically reactive with antibodies
15 and/or T cells generated against a polypeptide of the invention, particularly a
polypeptide having amino acid sequences disclosed herein, or to active fragments, or to
variants or biological functional equivalents thereof.

Likewise, a polypeptide composition of the present invention is
understood to comprise one or more polypeptides that are capable of eliciting antibodies
20 or T cells that are immunologically reactive with one or more polypeptides encoded by
one or more contiguous nucleic acid sequences contained in the amino acid sequences
disclosed herein, or to active fragments, or to variants thereof, or to one or more nucleic
acid sequences which hybridize to one or more of these sequences under conditions of
moderate to high stringency. Particularly illustrative polypeptides comprise the amino
25 acid sequence disclosed in SEQ ID NO: 2, 3, 5, 6, 7, 10-12, 14-15, 17-18, 20-23, 25-33,
39-47, 50-51, 54-64, 74-75, 90-97, 120-121, 122-140, 142-143, 153-178, 181, 195-205,
211-212, 215-216, 227-239, 241, 243, 248-250, 253-254 and 255-267.

As used herein, an active fragment of a polypeptide includes a whole or a
portion of a polypeptide which is modified by conventional techniques, *e.g.*,
30 mutagenesis, or by addition, deletion, or substitution, but which active fragment

exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of an HSV antigen or a variant or
5 biological functional equivalent thereof, as described herein. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that
10 is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of an HSV protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have
15 been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-
20 247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins).
25 Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native HSV protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is
30 similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art,

such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound
5 antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As noted above, a composition may comprise a variant of a native HSV protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native HSV protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In
10 other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or
15 antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

20 Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A
25 "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydrophobic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity,
30 hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino

acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, 5 pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or 10 addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or 15 other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above 20 may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and 25 plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange 30 resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art.

5 Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly,
10 Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino
15 acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements
20 responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide
25 as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (see, for example, Stoute et al. *New Engl. J. Med.*, 336:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is
30 derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises

approximately the first third of the protein (e.g., the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to
5 increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

10 In another embodiment, a *Mycobacterium tuberculosis*-derived Ra12 polynucleotide is linked to at least an immunogenic portion of an HSV polynucleotide of this invention. Ra12 compositions and methods for their use in enhancing expression of heterologous polynucleotide sequences is described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety.
15 Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been disclosed (U.S. Patent Application 60/158,585; see also, Skeiky *et al.*, *Infection and*
20 *Immun.* (1999) 67:3998-4007, incorporated herein by reference). The Ra12 C-terminal fragment of the MTB32A coding sequence expresses at high levels on its own and remains as a soluble protein throughout the purification process. Moreover, the presence of Ra12 polypeptide fragments in a fusion polypeptide may enhance the immunogenicity of the heterologous antigenic HSV polypeptides with which Ra12 is
25 fused. In one embodiment, the Ra12 polypeptide sequence present in a fusion polypeptide with an HSV antigen comprises some or all of amino acid residues 192 to 323 of MTB32A.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is
30 derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986).

LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for
5 expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

10 In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at
15 least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and
20 antigen-binding fragments thereof, that specifically bind to a HSV protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a HSV protein if it reacts at a detectable level (within, for example, an ELISA) with a HSV protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two
25 separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex

formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without HSV infection using the representative assays provided herein. For example, preferably, antibodies or other binding agents that bind to a HSV protein will generate a signal indicating the presence of infection in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without an HSV infection. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or biopsies) from patients with and without HSV (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. See, *e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen

is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a
5 suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the
10 desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells
15 and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture
20 supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable
25 vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

30 Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be

prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated
5 by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include
10 methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, *Shigella* toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a
15 suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group
20 containing a good leaving group (*e.g.*, a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or
25 an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in
30 the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups,

sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For

example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous,
5 intramuscular, subcutaneous and the like. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T
10 cells specific for HSV protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S.
15 Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a HSV polypeptide, polynucleotide encoding a HSV polypeptide and/or an antigen presenting cell (APC) that expresses
20 such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. In certain embodiments, HSV polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a HSV polypeptide if the T cells
25 specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such
30 assays may be performed, for example, as described in Chen et al., *Cancer Res.*

54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a HSV polypeptide (100 ng/ml - 100 μ g/ml, preferably 200 ng/ml - 25 μ g/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN- γ) is indicative of T cell activation (see Coligan et al., Current Protocols in Immunology, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a HSV polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. HSV protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a HSV polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a HSV polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a HSV polypeptide. Alternatively, one or more T cells that proliferate in the presence of a HSV protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz et al., 1997; Hwang et al., 1998; U.S. Patent No. 5,641,515; U.S. Patent No. 5,580,579 and U.S. Patent No. 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as

magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as
5 coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup of elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be
10 pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of
15 course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration,
20 product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash,
25 dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed
30 in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants.

Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U.S. Patent No. 5,543,158; U.S. Patent No. 5,641,515 and U.S. Patent No. 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U.S. Patent No. 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (e.g., glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought

about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered
5 isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml
10 of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human
15 administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally,
20 dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active
25 ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such
30 as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can

also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount
5 as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use
10 of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities
15 and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be
20 prepared. The preparation can also be emulsified.

3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the
25 lungs *via* nasal aerosol sprays has been described *e.g.*, in U.S. Patent No. 5,756,353 and U.S. Patent No. 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga et al., 1998) and lysophosphatidyl-glycerol compounds (U.S. Patent No. 5,725,871, specifically incorporated herein by reference in its entirety) are also well-
30 known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of

a polytetrafluoroethylene support matrix is described in U.S. Patent No. 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur et al., 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U.S. Patent No. 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran et al., 1997; Margalit, 1995; U.S. Patent No. 5,567,434; U.S. Patent No. 5,552,157; U.S. Patent No. 5,565,213; U.S. Patent No. 5,738,868 and U.S. Patent No. 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen et al., 1990; Muller et al., 1990). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath et al., 1986; Balazsovits et al., 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul et al., 1987), enzymes (Imaizumi et al., 1990a; Imaizumi et al., 1990b), viruses (Faller and Baltimore, 1984), transcription

factors and allosteric effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In addition, several successful clinical trails examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein et al., 1985a; 1985b; Coune, 1988; Sculier et al., 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)). MLVs generally have diameters of from 25 nm to 4 μ m. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.*, in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur et al. (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

Liposomes interact with cells via four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range

from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

10 Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface
15 components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable
20 nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland et al., 1987; Quintanar-Guerrero et al., 1998; Douglas et al., 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable
25 polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be easily made, as described (Couvreur et al., 1980; 1988; zur Muhlen et al., 1998; Zambaux et al. 1998; Pinto-Alphandry et al., 1995 and U.S. Patent No. 5,145,684, specifically incorporated herein by reference in its entirety).

VACCINES

In certain preferred embodiments of the present invention, vaccines are provided. The vaccines will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant.

5 An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes (into which the compound is incorporated; see, *e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example,

10 M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other HSV antigens may be present, either incorporated into a fusion polypeptide or as a

15 separate compound, within the composition or vaccine.

Illustrative vaccines may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems,

20 bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve

25 the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are

30 disclosed, for example, in Fisher-Hoch et al., *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine*

8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler et al., *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that a vaccine may comprise both a polynucleotide and a polypeptide component. Such vaccines may provide for an enhanced immune response.

It will be apparent that a vaccine may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (e.g., salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (e.g., sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the vaccine compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for

example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. Modified hepatitis B core protein carrier systems are also suitable, such as those described in WO/99 40934, and references cited therein, all incorporated herein by reference. One may also employ a carrier comprising the
5 particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextran), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants,
10 bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known
15 technology.

Any of a variety of immunostimulants may be employed in the vaccines of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A,
20 *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of
25 calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is
30 preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (e.g., IFN- γ , TNF α , IL-2 and IL-12) tend to favor the

induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

10 Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; see US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in
15 which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato et al., *Science* 273:352, 1996. Another preferred adjuvant is a saponin, preferably QS21
20 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO
25 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS
30 series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT)

and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties. Other preferred adjuvants comprise polyoxyethylene ethers, such as those described in WO 5 99/52549A1.

Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or 10 gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (see, *e.g.*, Coombes et al., *Vaccine* 14:1429-1438, 1996) and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a 15 polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co- 20 glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (see, *e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and 25 WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific 30 immune response that targets HSV-infected cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and

other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-HSV effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA
5 haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to
10 be effective as a physiological adjuvant for eliciting prophylactic or therapeutic immunity (see Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T
15 cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (see Zitvogel et al., *Nature Med.* 4:594-600,
20 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of
25 monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

30 Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized

phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature
5 phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a
10 HSV protein (or portion or other variant thereof) such that the HSV polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be
15 administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi et al., *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or
20 progenitor cells with the HSV polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated
25 immunological partner, separately or in the presence of the polypeptide.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or
30 aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be

stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

IMMUNOTHERAPEUTIC APPLICATIONS

In further aspects of the present invention, the compositions described
5 herein may be used for immunotherapy of HSV infections. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. The above pharmaceutical compositions and vaccines may be used to prophylactically prevent or ameliorate the extent of infection by HSV or to treat a patient already
10 infected with HSV. Administration may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical, and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous
15 host immune system to react against HSV infection with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established
20 HSV-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate therapeutic effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells
25 and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and
30 in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture
5 conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage,
10 monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy
15 must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (see, for example, Cheever et al., *Immunological Reviews* 157:177, 1997).

Alternatively, a vector expressing a polypeptide recited herein may be
20 introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary or intraperitoneal.

Routes and frequency of administration of the therapeutic compositions
25 described herein, as well as dosage, will vary from individual to individual, but may be readily established using standard techniques. In one embodiment, between 1 and about 10 doses may be administered over a 52 week period. In another embodiment, about 6 doses are administered, at intervals of about 1 month, and booster vaccinations are typically be given periodically thereafter. Alternate protocols may be appropriate for
30 individual patients.

A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-HSV immune response, and is preferably at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored, for example, by measuring the anti-HSV antibodies in a patient. Such
5 vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 µg
10 to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical
15 outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a HSV protein may correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples
20 obtained from a patient before and after treatment.

HSV DETECTION AND DIAGNOSIS

In general, HSV may be detected in a patient based on the presence of one or more HSV proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or other appropriate tissue) obtained
25 from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of HSV in a patient. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a HSV protein, which is also indicative of the presence or absence of HSV infection.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. See, *e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of HSV in a patient may be determined by
5 contacting a biological sample obtained from a patient with a binding agent and detecting in the sample a level of polypeptide that binds to the binding agent.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection
10 reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a
15 polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length HSV
20 proteins and portions thereof to which the binding agent binds, as described above.

The solid support may be any material known to those of ordinary skill in the art to which the protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a
25 plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention,
30 the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the agent and

functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time
5 varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

10 Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an
15 aldehyde group on the support with an amine and an active hydrogen on the binding partner (see, *e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized
20 on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of
25 detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as
30 bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to

bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with an HSV infection.

- 5 Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is
10 generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

- 15 The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed
20 for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme
25 reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

- To determine the presence or absence of HSV, the signal detected from the reporter group that remains bound to the solid support is generally compared to a
30 signal that corresponds to a predetermined cut-off value. In one embodiment, the cut-off value for the detection of HSV is the average mean signal obtained when the

immobilized antibody is incubated with samples from patients without HSV. In an alternate embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, 5 the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive. 10

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of 15 bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the 20 presence of HSV. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a 25 positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding 30

fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

5 Of course, numerous other assay protocols exist that are suitable for use with the HSV proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use HSV polypeptides to detect antibodies that bind to such polypeptides in a biological
10 sample. The detection of such protein-specific antibodies can allow for the identification of HSV infection.

 HSV infection may also, or alternatively, be detected based on the presence of T cells that specifically react with a HSV protein in a biological sample. Within certain methods, a biological sample comprising $CD4^+$ and/or $CD8^+$ T cells
15 isolated from a patient is incubated with a HSV polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by
20 Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for about 2-9 days (typically about 4 days) at 37°C with polypeptide (e.g., 5 - 25 μ g/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of HSV polypeptide to serve as a control. For $CD4^+$ T cells, activation is preferably detected by evaluating proliferation of the T cells. For
25 $CD8^+$ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of HSV in the patient.

 As noted above, HSV infection may also, or alternatively, be detected
30 based on the level of mRNA encoding a HSV protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain

reaction (PCR) based assay to amplify a portion of a HSV cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the HSV protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a HSV protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the HSV protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a HSV protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (see, for example, Mullis et al., *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not infected with HSV. The amplification reaction may be performed on several dilutions of cDNA, for example spanning two orders of magnitude.

As noted above, to improve sensitivity, multiple HSV protein markers may be assayed within a given sample. It will be apparent that binding agents specific

for different HSV polypeptides may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of HSV protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for HSV proteins provided
5 herein may be combined with assays for other known HSV antigens.

The present invention further provides kits for use within any of the above diagnostic and/or therapeutic methods. Such kits typically comprise two or more components necessary for performing a diagnostic and/or therapeutic assay and will further comprise instructions for the use of said kit. Components may be compounds,
10 reagents, containers and/or equipment. For example, one container within a diagnostic kit may contain a monoclonal antibody or fragment thereof that specifically binds to a HSV protein. Such antibodies or fragments may be provided attached to a support material, as described above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively,
15 contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a HSV protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a
20 polynucleotide encoding a HSV protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a HSV protein.

EXAMPLES

25 The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLE 1

IDENTIFICATION OF HSV-2 ANTIGENS

The following examples are presented to illustrate certain embodiments of the present invention and to assist one of ordinary skill in making and using the same.

- 5 The examples are not intended in any way to otherwise limit the scope of the invention.

Source of HSV-2 positive donors: Lymphocytes were obtained from two types of donors: Group A) seropositive donors with unknown clinical status, and Group B) seropositive donors with well characterized clinical status (viral shedding and ano-genital lesion recurrences).

- 10 Group A: Blood samples (50 ml) were obtained from 13 potential donors. No information regarding clinical history of HSV-2 infection was requested. The blood was screened for serum antibody against HSV-1 and HSV-2 by Western blot. PBMCs were also screened for specific proliferative T cell responses to HSV-1 and HSV-2 lysate antigens (ABI; Columbia, MD). Three donors (AD104, AD116, and
15 AD120) were positive for HSV-2 serum antibody and their PBMCs specifically proliferated in response to HSV-2 antigen. Leukopheresis PBMC were collected from these donors and cryopreserved in liquid nitrogen.

- Group B: Ano-genital lesion biopsies were obtained from donors DK21318 and JR5032. Lesion biopsy lymphocytes were expanded in vitro with IL-2
20 and PHA in the presence of 50 uM acyclovir and subsequently cryopreserved in liquid nitrogen. Typically 5×10^6 to 5×10^7 lymphocytes are obtained after two weeks. Autologous PBMC were also collected from the blood of DK2318 and JR5032 and cryopreserved in liquid nitrogen.

- Generation of CD4+ T cell lines: Cryopreserved PBMCs or lesion-
25 biopsy lymphocytes were thawed and stimulated in vitro with 1 ug/ml HSV-2 antigen (ABI) in RPMI 1640 + 10% human serum + 10 ng/ml IL-7. Irradiated autologous PBMC were added as antigen presenting cells for the lesion biopsy lymphocytes only. Recombinant IL-2 (1 ng/ml) was added on days 1 and 4. The cells were harvested, washed, and replated in fresh medium containing IL-2 and IL-7 on day 7. Recombinant
30 IL-2 was again added on day 10. The T cells were harvested, washed, and restimulated in vitro with HSV-2 antigen plus irradiated autologous PBMC in the same manner on

day 14 of culture. The T cell lines were cryopreserved at 1×10^7 cells/vial in liquid nitrogen on day 11-12 of the secondary stimulation. After thawing, the cryopreserved T cells retained the ability to specifically proliferate to HSV-2 antigen in vitro. These T cells were subsequently used to screen HSV-2 gene-fragment expression cloning
5 libraries prepared in *E. coli*, as described below.

Preparation of HSV-2 (333) DNA: HSV-2 strain 333 virus was grown in Vero cells cultured in roller bottles in 200 ml/bottle of Medium 199 (Gibco) + 5% FCS. Vero cells are transformed African green monkey fibroblast-like cells that were obtained from ATCC (Cat. # CCL-81). Near-confluence Vero cells (10 roller bottles)
10 were infected with HSV-2 strain 333 virus at an MOI of 0.01 in 50 ml/bottle of Medium 199 + 1% FCS. Cells and medium were harvested from the roller bottles and the cells pelleted. The supernatant was saved on ice and the cell pellets were resuspended in fresh Medium 199 + 1% FCS and lysed by 6 cycles of freezing/thawing. The cell debris in the lysates was pelleted and the supernatant pooled with the saved culture
15 supernatant. Virus was pelleted from the pooled supernatants by ultracentrifugation (12,000g, 2 hours, 4°C) and resuspended in 2 ml of fresh Medium 199 + 1% FCS. The virus was further purified on a 5 – 15% linear Ficoll gradient by ultracentrifugation (19,000 g, 2 hours, 4°C) as previously described (Chapter 10: Herpes simplex virus vectors of Molecular Virology: A Practical Approach (1993); Authors: F.J. Rixon and J.
20 McClaughlan, Editors: A.J. Davison and R.M. Elliott; Publisher: Oxford University Press, Inc, New York, N.Y.). The HSV-2 virus-containing band was extracted from the gradient, diluted 10-fold with Medium 199, and the virus pelleted by ultracentrifugation at 19,000 g for 4 hours at 4°C. The virus pellet was recovered and resuspended in 10 ml of Tris/EDTA (TE) buffer. Intact virions were treated with DNase and RNase to
25 remove cellular DNA and RNA. The enzymes were then inactivated by addition of EDTA and incubation at 65°C. DNA was prepared from the gradient-purified virus by lysis of the viral particles with SDS in the presence of EDTA, followed by phenol/chloroform extraction to purify the genomic viral DNA. HSV-2 DNA was precipitated with EtOH and the DNA pellet was dried and resuspended in 1 ml of
30 Tris/EDTA buffer. The concentration and purity of the DNA was determined by reading the OD 260 and OD 280 on a UV spectrophotometer. Genomic DNA prepared

in this manner was used for construction of an HSV-2 genomic fragment expression library in *E. coli*.

Construction of HSV-2 DNA fragment libraries in the pET17b vector:

The HSV2-I library was constructed as follows. DNA fragments were generated by
5 sonicating genomic HSV-2 DNA for 4 seconds at 15% output with a Fisher "60
SonicDismembrator" (Fisher). The sonicated DNA was then precipitated, pelleted, and
resuspended in 11 uL TE buffer. The approximate size of the DNA fragments was
measured by agarose gel electrophoresis of 1 uL of the fragmented HSV-2 genomic
DNA vs. 1.5 ug unsonicated material. The average size of the DNA fragments was
10 determined to be approx. 500 bp when visualized after ethidium bromide staining of the
gel. Incomplete DNA fragment ends were filled in (blunted) using T4 DNA polymerase.
EcoR1 adapters were then ligated to the blunt ends of the DNA fragments using T4
DNA ligase. The DNA was then kinased using T4 Polynucleotide Kinase, purified
using a manually loaded column of S-400-HR Sephacryl (Sigma) and ligated into the
15 pET17b expression vector. The HSV2-II library was constructed in a similar fashion.
The average size of inserts in this library was determined to be approximately 1000bp.

Generation of the HSV-2 fragment expression library in *E. coli*. The
HSV2-I library was transformed into *E. coli* for preparation of glycerol stocks and
testing of HSV-2 DNA insert representation. The DNA was transformed into
20 ElectroMAX DH10B *E. coli* (Gibco) in order to prepare a large quantity of HSV-
2/pET17b library DNA. Transformed bacteria were grown up on 3 LB/Ampicillin
plates (approx. 750 CFU/plate), a small subset of colonies were picked for sequencing
of DNA inserts, and the remaining bacteria from each plate collected as a pool for
preparation of plasmid DNA. These pools were named HSV-2 Pools 9, 10 and 11.
25 Glycerol stocks of a portion of these bacterial pools were stored at -80°C. Plasmids
were purified from the remainder of the pools. Equal quantities of plasmid DNA from
each of the 3 pools was combined to make a single pool of plasmid DNA. The
transformation efficiency of the pooled DNA was empirically determined using
JM109(DE3) *E. coli* bacteria. JM109(DE3) bacteria were then transformed with an
30 amount of the final pool of library DNA that was expected to yield 15 colony-forming
units (CFU) per plate. The transformed bacteria were then plated on 100 LB/amp

plates. Twenty CFU (on average) were actually observed on each of the 100 plates; therefore the pool size of this HSV-2 library was about 20 clones/pool. The bacterial colonies were collected as a pool from each plate in approximately 800 ul/plate of LB + 20% glycerol. Each pool was distributed equally (200 ul/well) among four 96-well U-bottom plates and these "master stock" plates were stored at -80°C. The size of this HSV-2 gene-fragment library (hereafter referred to as HSV2I) was therefore 96 pools of 20 clones/pool. Plasmid DNA was prepared from 20 randomly picked colonies and the inserts sequenced. Approximately 15% (3/20) contained HSV-2 DNA as insert, 80% (16/20) contained non-HSV-2 DNA (E. coli or Vero cell DNA), and 5% (1/20) contained no insert DNA. The HSV2-II DNA library was transformed into E. coli and random colonies analyzed in a similar manner. Relevant differences in the construction of library HSV2-II included the transformation of the HSV-2/pET17b ligation product into NovaBlue (Novagen) chemically competent E. coli instead of using electroporation for preparation of a larger quantity of plasmid for pooling and transformation into JM109(DE3) bacteria for empirical evaluation. Additionally, plasmid DNA was prepared from 10 pools averaging 160 colonies/plate. These 10 plasmid pools were combined in an equivalent fashion (normalized based on spectrophotometer readings) into one pool for transformation into JM109(DE3) as per previously, yielding an average of 20 colonies(clones)/plate for harvesting into glycerol stock pools as before. Approximately 25% contained HSV-2 DNA as insert, with the remaining 75% containing E. coli DNA as insert.

Induction of the HSV-2 fragment expression library for screening with human CD4+ T cells. One of the master HSV2I library 96-well plates was thawed at room temperature. An aliquot (20 uL) was transferred from each well to a new 96 well plate containing 180 uL/well of LB medium + ampicillin. The bacteria were grown up overnight and then 40 ul transferred into two new 96-well plates containing 160 uL 2xYT medium + ampicillin. The bacteria were grown for 1 hr.15 min at 37°C. Protein expression was then induced by addition of IPTG to 200 mM. The bacteria were cultured for an additional 3 hrs. One of these plates was used to obtain spectrophotometer readings to normalize bacterial numbers/well. The second, normalized plate was used for screening with CD4+ T cells after pelleting the bacteria

(approx. 2×10^7 /well) and removing the supernatants. The HSV2-II library was grown and induced in a similar fashion.

Preparation of autologous dendritic APC's: Dendritic cells (DCs) were generated by culture of plastic-adherent donor cells (derived from 1×10^8 PBMC) in 6 well plates (Costar 3506) in RPMI 1640 + 10% of a 1:1 mix of FCS:HS + 10 ng/ml GM-CSF + 10 ng/ml IL-4 at 37°C. Non-adherent DCs were collected from plates on day 6 of culture and irradiated with 3300 Rads. The DCs were then plated at 1×10^4 /well in flat-bottom 96-well plates (Costar 3596) and cultured overnight at 37°C. The following day, the DCs were pulsed with the induced HSV2-I or HSV2-II library pools by resuspending the bacterial pellets in 200 ul RPMI 1640 + 10 %FCS without antibiotics and transferring 10 ul/well to the wells containing the DCs in 190ul of the same medium without antibiotics. The DCs and bacteria were co-cultured for 90 minutes at 37°C. The DCs were then washed and resuspended in 100 ul/well RPMI 1640 + 10% HS + L-glut. + 50 ug/ml gentamicin antibiotic.

Preparation of responder T cells: Cryopreserved CD4+ T cell lines were thawed 5 days before use and cultured at 37°C in RPMI 1640 + 10% HS + 1 ng/ml IL-2 + 10 ng/ml IL-7. After 2 days, the medium was replaced with fresh medium without IL-2 and IL-7.

Primary screening of the HSV-2 libraries: The T cells were resuspended in fresh RPMI 1640 + 10% HS and added at 2×10^4 /well to the plates containing the E. coli-pulsed autologous DC's. After 3 days, 100 ul/well of supernatant was removed and transferred to new 96 well plates. Half of the supernatant was subsequently tested for IFN-gamma content by ELISA and the remainder was stored at -20°C. The T cells were then pulsed with 1 uCi/well of [3H]-Thymidine (Amersham/Pharmacia; Piscataway, NJ) for about 8 hours at 37°C. The 3H-pulsed cells were then harvested onto UniFilter GF/C plates (Packard; Downers Grove, IL) and the CPM of [3H]-incorporated subsequently measured using a scintillation counter (Top-Count; Packard). ELISA assays were performed on cell supernatants following a standard cytokine-capture ELISA protocol for human IFN-g.

From the HSV2-I library screening with T cells from D104, wells HSV2I_H10 and HSV2I_H12, for which both CPM and IFN-g levels were significantly above background, were scored as positive.

Breakdown of positive HSV2I library pools: The positive wells
5 (HSV2I_H10 and HSV2I_H12) from the initial CD4+ T cell screening experiment were grown up again from the master glycerol stock plate. Forty-eight sub-clones from each pool were randomly picked, grown up and IPTG-induced as described previously. The subclones were screened against the AD104 CD4+ T cell line as described above. A clone (HSV2I_H12A12) from the HSV2I_H12 pool breakdown scored positive. This
10 positive result was verified in a second AD104 CD4+ T cell assay.

Identification of UL39 as a CD4+ T cell antigen: The positive clone (HSV2I_H12A12) was subcloned and 10 clones picked for restriction digest analysis with EcoRI NB#675 pg. 34. All 10 clones contained DNA insert of the same size (approximately 900 bp in length). Three of these clones (HSV2I_H12A12_1, 7, and 8)
15 were chosen for sequencing and all contained identical insert sequences at both the 5' and 3' ends of the inserts. The DNA sequence of the insert is set forth in SEQ ID NO:1, and contains an open reading frame set forth in SEQ ID NO:2. The insert sequence was compared to the complete genomic sequence of HSV-2 strain HG52 (NCBI site, Accession #Z86099) and the sequence was determined have a high degree of homology
20 to UL39 (a.k.a. ICP6), the large subunit (140 kD) of the HSV ribonucleotide reductase, the sequence of which is set forth in SEQ ID NO:3. The insert sequence set forth in SEQ ID NO: 1 spans nucleotides 876 – 1690 of the UL39 open reading frame (3,432 bp) and encodes the amino acid sequence set forth in SEQ ID NO:2, which has a high degree of homology to amino acids 292 – 563 of UL39 (full length = 1143 aa).

25 Identification of US8A, US3/US4, UL15, UL18, UL27 and UL46 as CD4+ T cell antigens: In a manner essentially identical to that described above for the identification of UL39 as a T cell antigen, an additional HSV-2 gene fragment expression cloning library, referred to as HSV2-II, was prepared, expressed in E. coli, and screened with donor T cells.

30 Screening the HSV2-II library with T cells from donor AD116 identified the clone HSV2II_US8AfragD6.B_B11_T7Trc.seq, determined to have an insert

sequence set forth in SEQ ID NO:4, encoding open reading frames having amino acid sequences set forth in SEQ ID NO:5 and 6, with the sequence of SEQ ID NO:5 having a high degree of homology with the HSV-2 US8A protein, the sequence of which is set forth in SEQ ID NO:7.

5 In addition, screening the HSV2-II library with T cells from donor AD104 identified the following clone inserts:

SEQ ID NO:8, corresponding to clone HSV2II_US3/US4 fragF10B3_T7Trc.seq, containing a potential open reading frame having an amino acid sequence set forth in SEQ ID NO: 10;

10 SEQ ID NO:9, corresponding to clone HSV2II_US3/US4 fragF10B3_T7P.seq, containing an open reading frame having an amino acid sequence set forth in SEQ ID NO: 11, sharing a high degree of homology with the HSV-2 US3 protein (SEQ ID NO: 12);

SEQ ID NO:13, corresponding to clone
15 HSV2II_UL46fragF11F5_T7Trc.seq, containing an open reading frame having an amino acid sequence set forth in SEQ ID NO: 14, sharing a high degree of homology with the HSV-2 UL46 protein (SEQ ID NO: 15);

SEQ ID NO:16, corresponding to clone HSV2II_UL27frag-
H2C7_T7Trc.seq, containing an open reading frame having an amino acid sequence set
20 forth in SEQ ID NO:17, sharing a high degree of homology with the HSV-2 UL27 protein (SEQ ID NO:18);

SEQ ID NO:19, corresponding to clone
HSV2II_UL18fragF10A1_rc.seq, containing open reading frames having amino acid
sequences set forth in SEQ ID NO:20, 21 and 22, with SEQ ID NO:22 sharing a high
25 degree of homology with the HSV-2 UL18 protein (SEQ ID NO: 23); and

SEQ ID NO:24, corresponding to clone
HSV2II_UL15fragF10A12_rc.seq, containing an open reading frame having an amino
acid sequence set forth in SEQ ID NO: 25, sharing a high degree of homology with the
HSV-2 UL15 protein (SEQ ID NO: 26).

EXAMPLE 2

IDENTIFICATION OF HSV-2 ANTIGENS

CD4⁺ T cells from AD104 were found to recognize inserts from clones HSV2II_UL46fragF11F5_T7Trc.seq (SEQ ID NO: 13) and
5 HSV2II_UL18fagaF10A1_rc.seq (SEQ ID NO: 19) as described in detail in Example 1. The sequences from these clones share a high degree of homology to the HSV2-I genes, UL46 (SEQ ID NO: 15) and UL18 (SEQ ID NO:23), respectively. Therefore to further characterize the epitopes recognized by these T cells, overlapping 15-mer peptides were made across the clone insert fragments of UL18 and UL46. Peptide recognition by
10 AD104's CD4⁺ T cells was tested in a 48 hour IFN- γ ELISPOT assay. ELISPOTS were performed by adding 1×10^4 autologous EBV-transformed B cells (LCL) or DCs per well in 96 well ELISPOT plates. 2×10^4 AD104 CD4⁺ T cells from AD104's line were added per well with 5 μ g/ml of the HSV-2 peptides. AD104 CD4⁺ T cells recognized peptides 20 and 21 (SEQ ID NO: 32 and 33) of UL18, and peptides 1, 4, 9,
15 10, and 20 of UL46 (SEQ ID NO: 27-31).

EXAMPLE 3

IDENTIFICATION OF HSV-2 ANTIGENS

CD4⁺ T cell lines were generated from DK2318 and JR5032 lesion-biopsy. The CD4⁺ lymphocytes were stimulated twice in vitro on irradiated autologous
20 PBMC and HSV-2 antigen as described in example 1. The lines were tested for their antigen specificity as described in example 1 and cryopreserved. The CD4⁺ T cell lines were screened against the HSV2-II expression-cloning library generated in Example 1.

DK2318 was shown to react with clones C12 and G10. Clone C12 was determined to have an insert sequence set forth in SEQ ID NO:36. This insert was
25 found to have sequence homology with fragments of 2 HSV-II genes, nucleotides 723-1311 of UL23 and nucleotides 1-852 of UL22. These sequences correspond to amino acids 241-376 of UL23 as set forth in SEQ ID NO:40 and amino acids 1-284 as set forth in SEQ ID NO:41. The DNA sequence of SEQ ID NO:36 was searched against public databases including Genbank and shown to have a high degree of sequence homology to
30 the HSV-2 genes UL23 and UL22 set forth in SEQ ID NO:37 and 38 respectively. The

protein sequences encoded by SEQ ID NO:37 and 38 are set forth in SEQ ID NO:39 and 45. Clone G10 was determined to have an insert sequence which is set forth in SEQ ID NO:48, encoding open reading frames having an amino acid sequence set forth in SEQ ID NO:50, with the sequence of SEQ ID NO:48 having a high degree of
5 sequence homology with HSV-2 UL37, the sequence of which is set forth in SEQ ID NO:49, encoding open reading frames having the amino acid sequences set forth in SEQ ID NO:51. DK2318's CD4+ T cell line was screened against overlapping 15 mers covering the UL23 protein. DK2318's CD4 line was shown to react against three UL23 specific peptides (SEQ ID NO:41-43) suggesting that UL23 is a target.

10 The CD4+ T cell line generated from JR5032 was found to react with clone E9 which contained an insert sequence set forth in SEQ ID NO: 34, encoding open reading frames having amino acid sequences set forth in SEQ ID NO: 46, with SEQ ID NO: 34 having a high degree of sequence homology with HSV-2 RL2 (also referred to as ICP0), the sequence of which is set forth in SEQ ID NO:35, encoding an
15 open reading frame having the amino acid sequences set forth in SEQ ID NO:47.

EXAMPLE 4

CHARACTERIZATION OF CD4 CLONES F11F5 AND G10A9

Examples 2 and 3 describe the generation of CD4 T cell lines from donors AD104 and DK2313 which were screened against cDNA libraries generated
20 using the HSV-2 333 strain. AD104 was found to react against the clone HSV2II_UL46fragF11F5. This insert was partially sequenced with the sequence being disclosed in SEQ ID NO:13. Full length sequencing of the insert revealed that it encoded a fragment of UL46 which was derived from the HSV-2 333 strain. The DNA and amino acid sequences from this insert are disclosed in SEQ ID NO:52 and 54,
25 respectively.

DK2312 was found to react against the clone G10. This insert was partially sequenced and the sequence was disclosed in SEQ ID NO:48. Full length sequencing revealed that it encoded a fragment of UL37 which was derived from the HSV-2 333 strain. The DNA and amino acid sequences from this insert are disclosed in
30 SEQ ID NO:53 and 55, respectively.

EXAMPLE 5

IDENTIFICATION OF CD8-SPECIFIC IMMUNOREACTIVE PEPTIDES DERIVED FROM HSV-2

Peripheral blood mononuclear cells were obtained from the normal donors AD104, AD116, AD120, and D477. These donors were HLA typed using low-resolution DNA-typing methodology and the results are presented in Table 2.

TABLE 2

DONOR	AD104	AD116	AD120	D477
HLA-A	24, 33	0206, 24	0211, 3303	0201, 2501
HLA-B	45, 58	0702, 35	1505, 4403	1501, 5101
HLA-C	01, 0302	0702, 1203	0303, 0706	0304, 12

In order to determine which epitopes of HSV-2 were immunoreactive, synthetic peptides were synthesized. These peptides were 15 amino acids in length overlapping by 11 amino acids. The peptides were synthesized across the following regions of the following HSV-2 genes: UL15 (aa 600-734), UL18 (aa 1-110), UL23 (aa 241-376), UL46 (aa 617-722), US3 (aa 125-276), and US8A (aa 83-146).

CD8⁺ T cells were purified from the PBMC of each of the donors described above using negative selection. The purified CD8⁺ T cells were then tested for their reactivity against the HSV-2 specific peptides. Co-cultures containing 2x10⁵ CD8⁺ T cells, 1x10⁴ autologous dendritic cells and 10 µg/ml of a peptide pool (on average containing 10 peptides/pool) were established in 96 well ELISPOT plates that had been pre-coated with anti-human IFN-γ antibody (1D1K: mAbTech). After 24 hours, the ELISPOT plates were developed using a standard protocol well known to one of skill in the art. The number of spots per well were then counted using an automated video microscopy ELISPOT plate reader. CD8⁺ T cells from donors demonstrating a positive response against a peptide pool were then subsequently tested against the individual peptides in that pool in a second ELISPOT assay. The results of peptide reactivity are presented in Table 3.

TABLE 3

Donor	HSV-2 Gene	Peptide # (amino acid numbering)	SEQ ID NO
AD104	US3	#33 (262-276)	63
AD116	UL15	#23(688-702)	56
		#30(716-730)	57
	UL23	#7(265-279)	58
	UL46	#2(621-635)	59
		#8(645-659)	60
		#9(649-663)	61
		#11(657-671)	62
	US8A	#5(99-113)	64
AD120	UL46	Peptides: #1-12	-
D477	UL18	Peptides: #1-12	-
	UL23	Peptides: #1-20	-
	UL46	Peptides: #1-12	-

EXAMPLE 6

5 IDENTIFICATION OF HSV-2 ANTIGENS USING CD4⁺ T CELL CLONING

This Example describes the generation of CD4⁺ T cell clones from two donors. Donor JH is an HSV-2 seropositive donor who experiences infrequent recurrences of genital lesions and sheds virus infrequently, as determined by virus culture and PCR on daily swabs). HH is an HSV-2 exposed, but HSV-2 seronegative
10 donor.

CD4⁺ T cell clones for JH were generated by stimulating the donor's peripheral blood mononuclear cells (PBMC) for 14 days with UV-inactivated HSV-2, strain 333. Following two weeks of stimulation, the cells were cloned into 96 well plates using limiting dilution, and stimulated non-selectively using a monoclonal
15 antibody against CD3. Following 2 weeks of expansion, the clones were tested for their reactivity against UV-inactivated HSV-2, gB2 protein, gD2 protein and UL50. Clones 5 and 34 recognized gB2, clone 30 recognized gD2, and clone 11 recognized UL50.

Clones 39 and 47 were used for expression cloning. Antigen presenting cells (APC) used for both the expansion of the T cells and for the expression cloning were derived from HLA-matched normal donors. The clones were screened against two HSV-2 specific libraries, HSV2-II and HSV2-III.

5 Clone 39 was found to specifically recognize a partial sequence from UL39 presented by the HSV2-III library pools 1F4, 1G2, 2C4, and 3G11. The full length DNA sequence of UL39 is disclosed in SEQ ID NO:65, with the corresponding protein sequence disclosed in SEQ ID NO:3. The specific DNA sequence from pools 1F4, 1G2, and 3G11 that Clone 39 reacted against were identical. The inserts were
10 found to be 875 bp in length and the DNA sequence is disclosed in SEQ ID NO:66, with the corresponding amino acid sequence disclosed in SEQ ID NO:74. The insert from pool 2C4 was found to be 800 bp in length, the DNA sequence of which is disclosed in SEQ ID NO:67, with the corresponding amino acid sequence disclosed in SEQ ID NO:75.

15 Clone 47 was found to specifically recognize a partial sequence from ICP0 (RL2) presented by the HSV-2III library pools 2B2, 3A1, 3F12, 3H6, and 4B2. The full length DNA sequence of ICP0 was disclosed in SEQ ID NO:35, with the corresponding protein sequence disclosed in SEQ ID NO:47. The sequence inserts from pools 3H6, 3F12, and 4B2 were found to be identical, with an insert size of 1100 bp.
20 The DNA sequence corresponding to the 5' end of this sequence is disclosed in SEQ ID NO:68, with the 3' end disclosed in SEQ ID NO:69. The insert from pool 3A1 was found to be 1000bp in length, with the 5' portion of the DNA sequence disclosed in SEQ ID NO:70 and the 3' end of the insert disclosed in SEQ ID NO:71. The insert from pool 2B2 was found to be 1300 bp in length. The DNA sequence corresponding to
25 the 5' end of the insert is disclosed in SEQ ID NO:72, with the 3' end of the sequence disclosed in SEQ ID NO:73.

CD4⁺ T cell clones for HH were generated by stimulating the donor's peripheral blood mononuclear cells (PBMC) for 14 days with UV-inactivated HSV-2, strain 333. Following two weeks of stimulation, the cells were cloned into 96 well
30 plates using limiting dilution, and stimulated non-selectively using PHA. The clones were screened for their ability to proliferate in response to both HSV-1 and HSV-2

proteins. Clones 6, 18, 20, 22, 24, 27, 28, 29, 41, and 45 were all found to react strongly against HSV-1, however only clones 6, 18, 20, 22, and 24 were found to respond strongly to HSV-2. Therefore, clones 6, 18, 20, 22, and 24 were selected for expression cloning use. APC from an HLA-matched donor were used for *in vitro* expansion of the clones and for expression cloning. The clones were screened against two HSV-2 specific libraries, HSV2-II and HSV2-III (see Example 1 for details of libraries).

Clone 22 was found to recognize UL46 presented by the HSV2-II library, pools F7 and F11, in addition to pool 4E8 that was derived from the HSV2-III library.

10

EXAMPLE 7

GENERATION OF A UL19 EXPRESSING VACCINIA VIRUS

The UL19 gene was cloned into the Western Reserve Strain of Vaccinia Virus. This viral vector allows expression of UL19 in any cell infected with the vaccinia virus, or additionally, the vaccinia virus can be used to immunize humans or animals to generate immune responses against UL19.

In order to generate the vaccinia virus expressing UL19, the UL19 open reading frame (ORF), the sequence of which is disclosed in SEQ ID NO:76, was cloned from HSV-2 and inserted into the vaccinia virus shuttle plasmid, pSC11 (the DNA sequence of which is disclosed in SEQ ID NO:77). CV-1 cells transfected with the shuttle vector, pSC11/UL19, were co-infected with the wild-type Western Reserve Vaccinia Virus. In some cells, the shuttle plasmid underwent homologous recombination with the vaccinia virus, inserting the UL19 gene into the thymidine kinase location. These recombinant virions were isolated by plaque purification of 5-Bromo-deoxyuridine (BrdU) resistant virus that expressed Beta-galactosidase. The purified virus can then be used to infect cells to express the UL19 protein.

25

EXAMPLE 8

GENERATION OF A UL47 EXPRESSING VACCINIA VIRUS

The UL47 gene was cloned into the Western Reserve Strain of Vaccinia Virus. This viral vector allows expression of UL47 in any cell infected with the
5 vaccinia virus, or additionally, the vaccinia virus can be used to immunize humans or animals to generate immune responses against UL47.

In order to generate the vaccinia virus expressing UL47, the UL47 ORF, the sequence of which is disclosed in SEQ ID NO:78, was cloned from HSV-2 and inserted into the vaccinia virus shuttle plasmid, pSC11 (the DNA sequence of which is
10 disclosed in SEQ ID NO:77). CV-1 cells transfected with the shuttle vector, pSC11/UL47, were co-infected with the wild-type Western Reserve Vaccinia Virus. In some cells, the shuttle plasmid underwent homologous recombination with the vaccinia virus, inserting the UL47 gene into the thymidine kinase location. These recombinant virions were isolated by plaque purification of 5-Bromo-deoxyuridine (BrdU) resistant
15 virus that expressed Beta-galactosidase. The purified virus can then be used to infect cells to express the UL47 protein.

EXAMPLE 9

GENERATION OF A UL50 EXPRESSING VACCINIA VIRUS

The UL50 gene was cloned into the Western Reserve Strain of Vaccinia
20 Virus. This viral vector allows expression of UL50 in any cell infected with the vaccinia virus, or additionally, the vaccinia virus can be used to immunize humans or animals to generate immune responses against UL50.

In order to generate the vaccinia virus expressing UL50, the UL50 ORF, the sequence of which is disclosed in SEQ ID NO:79, was cloned from HSV-2 and
25 inserted into the vaccinia virus shuttle plasmid, pSC11 (the DNA sequence of which is disclosed in SEQ ID NO:77). CV-1 cells transfected with the shuttle vector, pSC11/UL50, were co-infected with the wild-type Western Reserve Vaccinia Virus. In some cells, the shuttle plasmid underwent homologous recombination with the vaccinia virus, inserting the UL50 gene into the thymidine kinase location. These recombinant
30 virions were isolated by plaque purification of 5-Bromo-deoxyuridine (BrdU) resistant

virus that expressed Beta-galactosidase. The purified virus can then be used to infect cells to express the UL50 protein.

EXAMPLE 10

GENERATION OF A UL49 EXPRESSING VACCINIA VIRUS

5 To facilitate intracellular degradation and Class I presentation of the Herpes Simplex Virus gene, UL49 (the DNA sequence of which is disclosed in SEQ ID NO:81), a fusion of the human Ubiquitin gene (the DNA sequence of which is disclosed in SEQ ID NO:80) and UL49 was constructed with the Ubiquitin gene located 5' of the UL49 gene. The last amino acid of the Ubiquitin ORF was mutated from glycine to
10 alanine to prevent co-translational cleavage of the fusion protein. After assembly of the fusion by PCR, it was cloned into the vaccinia virus shuttle vector, pSC11 (the DNA sequence of which is disclosed in SEQ ID NO:77). CV-1 cells transfected with the shuttle vector, pSC11/ubiquitin-UL49, were co-infected with the wild type Western Reserve Vaccinia Virus. In some cells the shuttle plasmid underwent homologous
15 recombination with the virus inserting the ubiquitin-UL49 gene into the thymidine kinase location. These recombinant virions were isolated by plaque purification of 5-Bromo-deoxyuridine (BrdU) resistant virus that expresses Beta-galactosidase. The purified virus can then be used to infect cells to express the UL49 protein.

The cells engineered to express UL49 are used to assay for specific
20 immune responses to UL49 protein. This vaccinia virus vector can also be used as a vaccine in humans to generate preventative or therapeutic responses against HSV-2.

EXAMPLE 11

EXPRESSION OF HERPES SIMPLEX VIRUS ANTIGENS IN E.COLI

This example describes the expression of recombinant HSV antigens
25 using an E. coli expression system combined with an N-terminal histadine tag.

Expression of HSV UL21 in E. coli:

The HSV UL21 coding region (the DNA sequence of which is disclosed in SEQ ID NO:85) was PCR amplified with the following primers:

PDM-602 5'gagctcagctatgccaccacc 3' (SEQ ID NO:98)

PDM-603 5'cggcgaattcattagtagaggcgggtgaaaaag3' (SEQ ID NO:99)

The PCR was performed with the following reaction components:

- 10 µl 10X Pfu buffer
- 5 1 µl 10 mM dNTPs
- 2 µl 10 µM of each primer
- 83 µl of sterile water
- 1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, Ca)
- 50 ng DNA

10 PCR amplification was performed using the following reaction conditions:

- 96°C for 2 minutes, followed by 40 cycles of:
- 96°C for 20 seconds;
- 60°C for 15 seconds; and
- 72°C for 2 minutes, followed by a final extension step of:
- 15 72°C for 4 minutes.

The PCR product was digested with EcoRI and cloned into pPDM His that had been cut with Eco72I and EcoRI. The amino acid sequence for the UL21-His construct was confirmed, and is disclosed in SEQ ID NO:91. The construct was then transformed into BLR pLys and BLR Codon Plus RP cells.

20 Expression of HSV UL39 in E. coli:

The HSV UL39 coding region (the DNA sequence of which is disclosed in SEQ ID NO:89) was PCR amplified from clone pET17b with the following primers:

PDM-466 5'cacgccgccgcacccaggcggac 3' (SEQ ID NO:100)

PDM-467 5'cggcgaattcattagtagaggcgggtgaaaaag 3' (SEQ ID

25 NO:101)

The PCR was performed with the following reaction components:

- 10 µl 10X Pfu buffer

- 1 μ l 10mM dNTPs
2 μ l 10 μ M of each primer
83 μ l of sterile water
1.5 μ l Pfu DNA polymerase (Stratagene, La Jolla, Ca)
5 50 ng DNA

PCR amplification was performed using the following reaction conditions:

- 96°C for 2 minutes, followed by 40 cycles of:
96°C for 20 seconds;
66°C for 15 seconds; and
10 72°C for 2 minutes, followed by a final extension step of:
72°C for 4 minutes.

- The PCR product was digested with EcoRI and cloned into pPDM His that had been cut with Eco72I and EcoRI. The amino acid sequence for the UL39-His construct was confirmed, and is disclosed in SEQ ID NO:90. The construct was then
15 transformed into BLR pLys and BLR Codon Plus RP cells.

Expression of HSV UL49 in E. coli:

The HSV UL49 coding region (the DNA sequence of which is disclosed in SEQ ID NO:83) was PCR amplified from clone pET17b with the following primers:

- PDM-466: 5'cacacctctcgccgctccgtcaagtc 3' (SEQ ID NO:102)
20 PDM-467: 5'cataagaattcactactcgagggggcgggcggggacg 3' (SEQ ID NO:103)

The PCR was performed with the following reaction components:

- 10 μ l 10X Pfu buffer
10 μ l 10X PCRx enhancer solution
25 3 μ l 10mM dNTPs
3 μ l 50mM mgSO₄
2 μ l 10 μ M of each primer
68 μ l of sterile water

1.0 µl Pfx polymerase (Gibco)

50 ng DNA

PCR amplification was performed using the following reaction conditions:

96°C for 2 minutes, followed by 40 cycles of:

5 96°C for 20 seconds;

67°C for 15 seconds; and

72°C for 2 minutes, followed by a final extension step of:

72°C for 4 minutes.

The PCR product was digested with EcoRI and cloned into pPDM His
10 that had been cut with Eco72I and EcoRI. The amino acid sequence for the UL49-His
construct was confirmed, and is disclosed in SEQ ID NO:97. The construct was then
transformed into BLR pLys and BLR Codon Plus RP cells.

Expression of HSV UL50 in E. coli:

The HSV UL50 coding region (the DNA sequence of which is disclosed
15 in SEQ ID NO:82) was PCR amplified from clone pET17b with the following primers:

PDM-458: 5'cacagtcagtgggggcccaggcgatcc 3' (SEQ ID NO:104)

PDM-459: 5'cctagaattcactagatgccagtggagccaaaccc 3' (SEQ ID NO:105)

The PCR was performed with the following reaction components:

10 µl 10X Pfu buffer

20 1 µl 10mM dNTPs

2 µl 10 µM of each primer

83 µl of sterile water

1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, Ca)

50 ng DNA

25 PCR amplification was performed using the following reaction conditions:

96°C for 2 minutes, followed by 40 cycles of:

96°C for 20 seconds;

68°C for 15 seconds; and
72°C for 2 minutes and 30 seconds, followed by a final extension step
of:

72°C for 4 minutes.

- 5 The PCR product was digested with EcoRI and cloned into pPDM His that had been cut with Eco72I and EcoRI. The amino acid sequence for the UL50-His construct was confirmed, and is disclosed in SEQ ID NO:96. The construct was then transformed into BLR pLys and BLR Codon Plus RP cells.

Expression of HSV UL19 in E. coli:

- 10 The HSV UL19 coding region (the DNA sequence of which is disclosed in SEQ ID NO:84) was PCR amplified from clone pET17b with the following primers:

PDM-453: 5'gccgctcctgcccgcgaccccc 3' (SEQ ID NO:106)

PDM-457: 5'ccagaattcattacagagacaggcccttagc 3' (SEQ ID NO:107)

The PCR was performed with the following reaction components:

- 15 10 µl 10X Pfu buffer
 1 µl 10mM dNTPs
 2 µl 10 µM of each primer
 83 µl of sterile water
 1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, Ca)
20 50 ng DNA

PCR amplification was performed using the following reaction conditions:

- 96°C for 2 minutes, followed by 40 cycles of:
 96°C for 20 seconds;
 70°C for 15 seconds; and
25 72°C for 4 minutes, followed by a final extension step of:
 72°C for 4 minutes.

The PCR product was digested with EcoRI and cloned into pPDM His that had been cut with Eco72I and EcoRI. The amino acid sequence for the UL19-His

construct was confirmed, and is disclosed in SEQ ID NO:95. The construct was then transformed into BLR pLys and BLR Codon Plus RP cells.

Expression of HSV UL47 in E. coli:

5 The HSV UL47 coding region (the DNA sequence of which is disclosed in SEQ ID NO:87) was PCR amplified using the following primers:

PDM-631: 5'cactccgtggcgcgggcatgccg 3' (SEQ ID NO:108)

PDM-632: 5'ccgtagaattcactatggcggtggcgggcc 3' (SEQ ID NO:109)

The PCR was performed with the following reaction components:

10 10 µl 10X Pfu buffer
1 µl 10mM dNTPs
2 µl 10 µM of each primer
83 µl of sterile water
1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, Ca)
50 ng DNA

15 PCR amplification was performed using the following reaction conditions:

96°C for 2 minutes, followed by 40 cycles of:
96°C for 20 seconds;
67°C for 15 seconds; and
72°C for 2 minutes and 30 seconds, followed by a final extension step
20 of:
72°C for 4 minutes.

The PCR product was digested with EcoRI and cloned into pPDM His that had been cut with Eco72I and EcoRI. The amino acid sequence for the UL47-His construct was confirmed, and is disclosed in SEQ ID NO:94. The construct was then
25 transformed into BLR pLys and BLR Codon Plus RP cells. Protein yields were low using this construct. UL47 was also cloned into pPDM Trx with two histadine tags that had been digested with StuI and EcoRI. The DNA and amino acid sequences for this

construct are disclosed in SEQ ID NOs:86 and 92, respectively. Protein yields were much higher using this fusion construct.

Four additional fragments of UL47, designated UL47 A-D were also PCR amplified.

5 The UL47 A coding region was amplified using the following primer pairs:

PDM-631: 5'cactccgtgcgcgggcatgccg 3' (SEQ ID NO:110)

PDM-645: 5'catagaattcatcacgcgcgggaggggctggttttgc 3' (SEQ ID NO:111)

The UL47 B coding region was amplified using the following primer pairs:

10 PDM-646: 5'gacacggtggtcgcgtgcgtggc 3' (SEQ ID NO:112)

PDM-632: 5'ccgttagaattcactatgggcgtggcgggcc 3' (SEQ ID NO:113).

Both fragments were amplified using the following PCR reaction components:

15 10 µl 10X Pfu buffer
1 µl 10mM dNTPs
2 µl 10 µM of each primer
83 µl of sterile water
1.5 µl Pfu DNA polymerase (Stratagene, La Jolla, Ca)
50 ng DNA

PCR amplification was performed using the following reaction conditions:

20 96°C for 2 minutes, followed by 40 cycles of:
96°C for 20 seconds;
67°C for 15 seconds; and
72°C for 2 minutes, followed by a final extension step of:
72°C for 4 minutes.

25 The UL47 C coding region was amplified using the following primer pairs:

PDM-631: 5'cactccgtgcgcgggcatgccg 3' (SEQ ID NO:114)

PDM-739: 5'cgatgaattcatcagacccacccgttg 3' (SEQ ID NO:115)

The UL47 D coding region was amplified using the following primer pairs:

PDM-740: 5'gtgctggcgacggggctcatcc3' (SEQ ID NO:116)

PDM-632: 5'ccgttagaattcactatggcggtggcgggcc 3' (SEQ ID NO:117).

5 Both fragments were amplified using the following PCR reaction components:

10 µl 10X Pfu buffer

1 µl 10mM dNTPs

2 µl 10 µM of each primer

83 µl of sterile water

10 1.5 µl Pfu DNA polymerase (Stratagene, LaJolla, Ca)

50 ng DNA

PCR amplification was performed using the following reaction conditions:

96°C for 2 minutes, followed by 40 cycles of:

96°C for 20 seconds;

15 63°C for 15 seconds; and

72°C for 2 minutes, followed by a final extension step of:

72°C for 4 minutes.

The PCR product from UL47 C was digested with EcoRI and cloned into pPDM His that had been digested with Eco72I and EcoRI. The sequence was confirmed then the construct was transformed into BLR pLys S and BLR CodonPlus RP cells. The DNA and amino acid sequences of UL47 C are disclosed in SEQ ID NOs:88 and 93, respectively.

EXAMPLE 12

IDENTIFICATION OF A NOVEL DNA SEQUENCE ENCODING THE HSV-2 GENE, US8

25 The US8 gene of HSV-2 was cloned from the laboratory HG52 viral strain and sequenced, the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:118 and 120, respectively. SEQ ID NO:118 was then compared to the

HSV-2 HG52 strain genomic sequence contained in GenBank (accession number Z86099), the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:119 and 121, respectively. This comparison revealed that SEQ ID NO:118 contained an extra base pair at position 542 that results in a frameshift. The presence of
5 the extra base pair was also confirmed in a second laboratory strain of HSV-2, 333. There was one additional base pair (bp 156) upstream of the first stop codon in SEQ ID NO:118 that differed from the GenBank US8 sequence (SEQ ID NO:119). No change in the US8 amino acid sequence would result from the change in the nucleotide sequence at base pair 156.

10 In addition to examining the sequence of a number of laboratory strains of HSV-2, genomic DNA sequence was also obtained from two clinically isolated viral samples, donors RW1874 and HV5101). Using PCR primers designed to gene specific sequences both up- and down-stream of the position 542 insertion, this region was PCR amplified and directly sequenced from the purified amplicon using the same primer
15 pair. The sequences obtained from both RW1874 and HV5101 showed the additional guanine nucleotide at position 542. HV5101 had one additional base pair change at base pair 571 (G/571/C:HV5101/location/HG52) when compared to HG52 (SEQ ID NO:119). This difference is a non-conservative change in the frameshift form.

20

EXAMPLE 13

VACCINATION WITH THE HSV-2 UL47 PROTEIN ELICITS BOTH A CD4⁺ AND CD8⁺ SPECIFIC T CELL RESPONSE

This example demonstrates the effectiveness of UL47 as a vaccine against HSV-2. Balb/c mice vaccinated with UL47, delivered by plasmid DNA,
25 mounted a UL47-specific CD4⁺ and CD8⁺ cell response.

Two Balb/c mice were immunized three times with 100µg of UL47 plasmid DNA (UL47 DNA), an additional four mice were immunized twice with UL47, followed by infection with 1x10³ pfu of an attenuated HSV-2 strain, 333vhsB (UL47 DNA/HSV). A further four mice received HSV-2 infection alone (HSV control). The

spleens were harvested two weeks post-final immunization and stimulated *in vitro* with vaccinia-UL47 for 7 days.

On day 7, the splenocytes were assayed for cytotoxic activity by chromium release against P815 cells pulsed with pools of 10-15-mer peptides that spanned the UL47 gene (18 pools total). The splenocytes were re-stimulated *in vitro* and then re-assayed against positive peptide pools, plus the constitutive 15-mer peptides. At an effector:target ratio of 100:1, specific lytic activity by CD8⁺ cells could be seen in response to P815 cells pulsed with peptides 85 (SEQ ID NO:122), 89 (SEQ ID NO:123), 99 and 98 (SEQ ID NO:124), 105 (SEQ ID NO:125), and 112 (SEQ ID NO:126).

In order to determine the presence of a CD4⁺ T cell responses, splenocytes were stimulated *in vitro* with 5µg/ml recombinant UL47 (rUL47). Three days following stimulation, the culture supernatants were harvested and assayed for IFN-gamma by ELISA. Supernatants harvested from both the splenocytes from the "UL47 DNA" mice (those that were immunized) and the "UL47 DNA/HSV" mice (those that were immunized followed by infection with HSV) had significant levels of IFN-gamma present compared to the "HSV control" mice (those who were unimmunized and infected).

A further four mice were immunized four times with UL47 DNA and their splenocytes harvested. The splenocytes were then stimulated with peptides p85, p89, p98, p99, p105, and p112 and the CD8⁺ cells assayed for the presence of intracellular IFN-gamma production using flow cytometry. The percentages of CD8⁺ cells producing IFN-gamma were significant in the splenocytes stimulated with peptides p85, p89, p98, p99, p105 and p112, compared to the control cells (cells stimulated with media or PBS alone). Responses seen against peptides p98 and p99 should be the highest percentages, with greater than 2% of all CD8⁺ splenocytes positive for intracellular IFN-gamma.

These data further demonstrate the effectiveness of UL47 as a vaccine candidate in the protection against or treatment of HSV infection.

EXAMPLE 14

CD8⁺ T CELL RESPONSES FROM HSV-2 SEROPOSITIVE DONORS

Six HSV-2 seropositive donors were screened to determine which HSV-2 proteins were capable of eliciting a CD8⁺ T cell response. The donors included: AD104, AD116, AD120, D477, HV5101, and JH6376. In order to determine which HSV-2 proteins were immunogenic, synthetic peptides (15-mers overlapping by 11 amino acids) were synthesized across the following region of several HSV-2 polypeptides, including: UL15 (a.a. 600-734), UL18 (a.a. 1-110), UL23 (a.a. 241-376), UL46 (a.a. 617-722), UL47 (a.a. 1-696), (a.a. 1-300), ICP27 (a.a. 1-512), US3 (a.a. 125-276), and US8A (a.a. 83-146). Peptides synthesized for UL47, UL49, and ICP27 spanned the full-length polypeptide. Peptides synthesized for UL15, UL18, UL23, UL46, US3, and US8A spanned the portions of these polypeptides previously determined to encode antigens recognized by CD4⁺ T cells during CD4 expression-cloning library screening.

The donors CD8⁺ T cells were isolated from PBMC using the following procedure: initially peripheral blood lymphocytes (PBL) were separate from macrophages using plastic adherence. The CD8⁺ T cells were then further purified by depletion of non-CD8⁺ cells using a commercial MACS bead kit (Miltenyi). CD8⁺ T cells isolated using this method are generally >95% CD8⁺/CD3⁺/CD4⁻, as measured by flow cytometry (FACS). Peptides were screened by 24-hour co-culture of CD8⁺ T cells (2x10⁵/well), autologous dendritic cells (1x10⁴/well), and peptides (10μg/ml each) in 96 well ELISPOT plates pre-coated with anti-human IFN-gamma antibody. Peptides were initially screened as pools of ≥10 peptides. ELISPOT plates were subsequently developed per a standard protocol. The numbers of spots per well were counted using an automated video-microscopy ELISPOT reader. Peptide from pools screening positive were subsequently tested individually in a second ELISPOT assay.

For AD104, only the peptide US3 #33 (SEQ ID NO:139: amino acids 262-276) scored positive.

For AD116, peptides UL15 #23 (SEQ ID NO:127: amino acids 688-702), UL15 #30 (SEQ ID NO:128: amino acids 716-730), UL23 #7 (SEQ ID NO:129: amino acids 265-272), UL46 #2 (SEQ ID NO:130: amino acids 621-635), UL46 #8

(SEQ ID NO:131: amino acids 645-659), UL46 #9 (SEQ ID NO:132: amino acids 649-663), UL46 #11 (SEQ ID NO:133: amino acids 657-671), UL47 #86 (SEQ ID NO:134: amino acids 341-355), UL49 #6 (SEQ ID NO:135: amino acids 21-35), UL49 #49 (SEQ ID NO:138: amino acids 193-208), and US8A #5 (SEQ ID NO:140: amino acids 99-113) scored positive both pooled and individually. In addition, AD116 also recognized the B*0702-restricted epitope UL49 #12 (SEQ ID NO:136: amino acids 45-59) and UL49 #13 (SEQ ID NO:137: amino acids 49 to 63).

Donors D477, HV5101, and JH6376 T cells recognized the HLA-A*0201-restricted epitopes UL47 #73/#74 (amino acids 289-297) and UL47 #137/#138 (amino acids 550-559), respectively.

Donor AD120 scored positive for one peptide pool, UL46 #1-12.

Donor D477 scored positive for 5 peptide pools: UL18 #1-12, UL23 #1-10, UL23 #11-20, UL46 #1-12, and UL49 #11-20.

EXAMPLE 15

IDENTIFICATION OF A NOVEL SEQUENCE CODING FOR THE US4 PROTEIN OF HSV-2

Screening the HSV2-II library with T cells from donor AD104 had previously identified the clone insert F10B3 (see Example 1 for details). SEQ ID NO:8, corresponds to the partial sequence of the insert from clone HSV2II_US3/US4 fragF10B3_T7Trc.seq, and contains a potential open reading frame having an amino acid sequence set forth in SEQ ID NO: 10. The full-length DNA and amino acid sequences corresponding to the insert sequence are disclosed in SEQ ID NOs:141 and 142, respectively. The full length US4 HG52 DNA and amino acid sequence are disclosed in SEQ ID NO:179 and 143, respectively, and differs from the insert sequence as follows: S35N (HG52/location/333).

EXAMPLE 16

IDENTIFICATION OF HSV-2-SPECIFIC CD8⁺ T CELL RESPONSES IN HSV-2

CD8⁺ T cells isolated from a panel of HSV-2 seropositive donors were screened for their ability to respond to a variety of HSV-2 proteins. Briefly, PBMCs were obtained from donors EB5491, AG10295, LM10295, and 447, and enriched for

CD8⁺ T cells using microbeads or CD8⁺ Enrichment Kits from Miltenyi. Synthetic peptides (15 amino acids in length and overlapping in sequence by 10 or 11 amino acids) were synthesized across several complete or partial ORFs from HSV-2 strain HG52, including proteins UL21 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:144 and 154, respectively), UL50 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:145 and 153, respectively), US3 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:146 and 154, respectively), UL54 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:147 and 156, respectively), US8 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:148 and 157, respectively), UL19 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:149 and 158, respectively), UL46 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:150 and 159, respectively), UL18 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:151 and 160, respectively), and RL2 (the full length DNA/amino acids of which are disclosed in SEQ ID NOs:152 and 161, respectively). The peptides were screened by 24 co-culture of the donor's CD8⁺ T cells ($2-5 \times 10^5$ cells/well), autologous dendritic cells ($2-5 \times 10^4$ cells/well) and peptides (0.5 µg/ml each) in 96-well ELISPOT plates that had been pre-coated with anti-human IFN-γ antibody. Each peptide pool was screened in an individual well. The ELISPOT plates were developed as per a standard protocol. The number of spots per well was counted using an automated video-microscopy ELISPOT reader. Individual 15-mer peptides, determined from peptide pools testing positive, were screened as described above and returned the following results:

Donor EB5491 demonstrated CD8⁺ T cell responses to the HSV-2 antigens: ICP0 peptide #43 (amino acids 211-225: IWTGNPRTAPRSLSL: SEQ ID NO:162). UL46 peptides #41 (amino acids 201-215: YMFFMRPADPSRPST: SEQ ID NO:163), UL46 #50 (amino acids 246-260: VCRRLGPADRRFVAL: SEQ ID NO:164), UL46 #51 (amino acids 251-265: GPADRRFVALSGSLE: SEQ ID NO:165), and UL46 #60 (amino acids 296-310: SDVLGHLTRLAHLWE: SEQ ID NO:166). Donor EB5491 also demonstrated a CD8⁺ T cell response to the HSV-2 protein, US8 #74 (amino acids 366-380: HGMTISTAAQYRNAV: SEQ ID NO:167).

Donor JH6376 demonstrated CD8+ T cells responses to the HSV-2 proteins ICP0, which corresponded to a 9-mer mapped to amino acids 215-223 (NPRTAPRSL: SEQ ID NO:177) and UL46, which corresponded to a 10-mer mapped to amino acids 251-260 (GPADRRFVAL: SEQ ID NO:178).

5 Donor AG1059 demonstrated CD8+ T cell responses to the HSV-2 proteins UL19 peptide 102 (amino acids 506-520: LNAWRQRLAHGRVRW: SEQ ID NO:168), UL19 #103 (amino acids 511-525: QRLAHGRVRWVAECQ: SEQ ID NO:169) and UL18 #17 (amino acids 65-79: LAYRRRFPAVITRVL: SEQ ID NO:172) and UL18 #18 (amino acids 69-83: RRFPVITRVLPTRI: SEQ ID NO:173).

10 Donor LM10295 demonstrated CD8+ T cell responses to the HSV-2 protein UL19 #74 (amino acids 366-380: DLVAIGDRLVFLEAL: SEQ ID NO:170) and UL19 #75 (amino acids 371-385: GDRLVFLEALERRIY: SEQ ID NO:171).

Donor 477 demonstrated CD8+ T cell responses to the HSV-2 protein UL50 #16 (amino acids 76-90: CAIHHAPAVSGPGPH: SEQ ID NO:174), UL50 #23 (amino acids 111-125: PNGTRGFAPGALRVD: SEQ ID NO:175), and UL50 #49 (amino acids 241-255: LRVLRAADGPEACYV: SEQ ID NO:176).

EXAMPLE 17

EXPRESSION OF A TRUNCATED FORM OF UL47 IN E. COLI

A C-terminal truncation of the full length UL47 coding region was expressed in E. coli, and designated as UL47F. This truncated portion of UL47 contains the C-terminal T cell epitope of UL47, corresponding to amino acids 500-559.

Expression of HSV UL47 F in E. coli:

The HSV UL47F coding region (the DNA and amino acid sequences of which are disclosed in SEQ ID NO:180 and 181, respectively) was PCR amplified using the following primers:

CBH-631: 5'ctgggtctggctgacacgggtgctgcgtgcgtg 3' (SEQ ID NO:182)

PDM-632: 5'ccgttagaattcactatgggcgtggcgggcc 3' (SEQ ID NO:183)

The PCR was performed with the following reaction components:

- 10 μ l 10X Pfu buffer
1 μ l 10mM dNTPs
2 μ l 10 μ M of each primer
5 83 μ l of sterile water
1.5 μ l Pfu DNA polymerase (Stratagene, La Jolla, Ca)
50 ng DNA

PCR amplification was performed using the following reaction conditions:

- 96°C for 2 minutes, followed by 40 cycles of:
10 96°C for 20 seconds;
68°C for 15 seconds; and
72°C for 1 minute and 30 seconds, followed by a final extension step of:
72°C for 4 minutes.

- The PCR product was digested with EcoRI and cloned into pPDM His
15 that had been cut with Eco72I and EcoRI. The sequence of the construct was
confirmed, and then the construct was transformed into BRL pLys S and BLR
CodonPlus RP cells.

EXAMPLE 18

IDENTIFICATION OF NOVEL ANTIGENS FROM HSV-2 RECOGNIZED BY

20 HUMAN CD8⁺ T CELLS

- This example illustrates the identification of multiple immunogenic
HSV-2 antigens using ELISPOT screening of HSV-2 peptides. These findings identify
HSV-2 antigens capable of eliciting a cellular immune response *in vivo*. Identification
of such antigens allows for the development of vaccines to protect against HSV-2
25 infection, as well as compounds that can be used in the treatment of HSV-2 infection.

A panel of HSV-2 seropositive donors including AD104, AD116, D574,
GI10897, RW1874, YS10063, AC10022, LB10802, NH9894, PA10939, VB10576,
MA11259 (seronegative), and JG10758 (described in detail in Table 4) were used to
identify immunogenic portions of HSV-2. Blood was obtained from each donor and

peripheral blood mononuclear cells (PBMCs) were isolated using a Ficoll gradient. The PBMCs were washed thoroughly with PBS/EDTA, and suspended in a 10% DMSO/50% FBS/40% RPMI solution and frozen.

5 TABLE 4

Donor	Serol. status	Shedding Status	HLA- A	HLA- B	HLA- Cw	HLA- DRB1	HLA- DQB1
AD104 413	1- 2+	Unknown	24 33	46 58	01 0302	04 09	03 04
AD116 421	1- 2+	Unknown	0206 24	0702 35	0702 1203	0408 1501	0304 0602
D574 146	1 2+	Unknown	11 68	27 55	02 03	0407 0901	0301 0303
RW1874	1- 2+	Frequent	0101 0201	0801 4501	0701 16	03 11	2 3
YS10063	1+ 2+	Infrequent (0)	02 33	35 58	03 04	0301 1401	0201 0503
AC10022	1- 2+	Frequent	02	27 37	01 06	0101 1104	0301 0501
JG10758	1+ 2+	Frequent	01 02	08 51		0101 0301	0201 0501
GI10897	1+ 2+		02 24	35		1104 1401	0301 0503
PA10939	1- 2+	Infrequent <i>Cult 2/70</i>	02 31	44 14	05 08	0101 0102	0501
VB10576	1- 2+	Infrequent <i>Cult 3/50</i>	24 26	35		1104 0407	0301
NH9894	1- 2+	Infrequent <i>Cult 2/46</i>	29 68	44 63	07 14	0401 1101	0301
LB10802	1+ 2+	Infrequent <i>Cult 0/26</i>	29 68	44 35	04 07	0401 1104	0301 0302
MA11259	1+ 2-	N/A	01 02	51 58	02 07	0701	0202

Synthetic peptides (15 amino acids in length and overlapping by 10 or 11 amino acids) were synthesized across several complete or partial open reading frames (ORFs) from the HG52 strain of HSV-2. These ORFs included UL18 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:184 and 195, respectively), LAT-ORF1 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:185 and 198, respectively), UL48 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:186 and 205, respectively), UL41 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:187 and 204, respectively), UL39 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:188 and 203, respectively), UL37 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:189 and 202, respectively), UL36 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:190 and 201, respectively), UL29 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:191 and 200, respectively), UL25 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:192 and 199, respectively), ICP4 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:193 and 197, respectively), ICP0 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:35 and 47, respectively), US3 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:146 and 154, respectively) and ICP22 (the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:194 and 196, respectively). Individual 15-mer peptide stocks were made by dissolving each peptide at a concentration of 10mg/ml in DMSO. Peptide pools containing between 16-90 peptides/pool were made from the individual peptide stocks by combining individual peptides at 100µg/ml.

The peptide pools were screened against CD8⁺ T cells enriched from the individual donors' PBMCs. CD8⁺ T cell enriching was performed using either CD8⁺ microbeads or CD8⁺ enrichment kits from Miltenyi, as per the manufacturer's instructions. The peptides were screened by 24-hour co-culture of CD8⁺ T cells (5x10⁵/well), autologous dendritic cells (5x10⁴/well) and peptide pools (0.5µg/ml each) in 96-well ELISPOT plates that had been pre-coated with anti-human IFN-γ antibody. Each peptide pool was screened in an individual well. ELISPOT plates were

subsequently developed per a standard protocol, and the number of spots per well counted using an automated video-microscopy ELISPOT reader.

The following donors responded to the following HSV-2 antigens and peptide pools:

5 Donor LB10802 responded to UL39, peptide pools C (amino acids 501-760) and D (amino acids 751-1142), UL21, peptide pools B (amino acids 260-530), UL19, peptide pools E (amino acids 1001-1374) and B (amino acids 251-510), and ICP0, pool C (amino acids 499-825).

Donor PA10939 responded to UL25, peptide pool B (amino acids 251-585), UL47, peptide pool C (amino acids 401-696) and UL46, peptide pool C (amino acids 500-722).

Donor NH9894 responded to UL39, peptide pool C (amino acids 501-760), UL36, peptide pool G (amino acids 2671-3122), UL29, peptide pool C (amino acids 501-760), UL25, peptide pool A (amino acids 1-260), UL49, amino acids 1-300, 15 UL47, peptide pool C (amino acids 401-696), UL46, peptide pool B (amino acids 251-510), UL19, peptide pool D (amino acids 751-1010), and ICP0, peptide pool A (amino acids 1-259).

Donor 574 responded UL39, peptide pool D (amino acids 751-1142), UL18 (amino acids 1-318), and ICP0, peptide pool B (amino acids 251-510).

20 Donor AD104 responded UL29, peptide pool C (amino acids 501-760), UL25, peptide pool A (amino acids 1-260), and ICP4, peptide pool D (amino acids 751-1010).

Donor YS10063 responded UL25, peptide pool B (amino acids 251-585), US3, amino acids 163-481, UL47 peptide pools C (amino acids 401-696) and B (amino acids 201-411), UL19, peptide pool C (amino acids 501-760) and ICP0, peptide pools C (amino acids 499-825) and A (amino acids 1-259).

Donor JG10758 responded to UL39, peptide pool C (amino acids 501-760).

Donor AD116 responded to UL41, peptide pool A (amino acids 1-260) 30 and ICP22, peptide pool B (amino acids 206-413).

Donor VB10576 responded to LAT-1, peptide pool B (amino acids 42-82), UL39, peptide pool C (amino acids 501-760), UL36, peptide pools A (amino acids 1-455) and C (amino acids 889-1345), ICP22, peptide pool B (amino acids 206-413), UL19, peptide pools C (amino acids 501-760) and A (amino acids 1-260), ICP27,
5 peptide pool B (amino acids 253-512), and ICP0, peptide pool C (amino acids 499-825).

Donor GI10897 responded to UL39, peptide pools C (amino acids 501-760) and A (amino acids 1-260), UL37, peptide pool C (amino acids 501-760), UL36, peptide pool C (amino acids 889-1345), and UL25, peptide pool B (amino acids 251-585).

10 Donor RW1874 responded to UL41, peptide pool B (amino acids 251-492), UL39, peptide pool C (amino acids 501-760), and UL25, peptide pool B (amino acids 251-585).

Donor AC10022 responded to UL18 amino acids 1-318, and ICP4, peptide pools D (amino acids 751-1010) and B (amino acids 251-510).

15 Donor MA11259 (HSV sero-negative donor) responded to UL39, peptide pool C (amino acids 501-760), ICP27, peptide pool B (amino acids 253-512), and ICP0, peptide pool A (amino acids 1-259).

EXAMPLE 19

IDENTIFICATION OF HSV-2 ANTIGENS USING CD4+ T CELL CLONING

20 Donor HH is a HSV-2 exposed, but uninfected, seronegative donor. The generation of a HSV-2 specific CD4+ T cell line was previously described in Example 6. This example illustrates the identification of immunoreactive HSV-2 antigens using T cell expression cloning of E. coli gene-fragment expression libraries. These libraries were generated using the 333 strain of HSV-2. These experiments were performed
25 essentially as described in Example 11.

Three distinct library inserts were identified from the HSV-2 library:

Inserts 1/A7 and 1/F3 span the corresponding HSV-2 (strain HG52) genome at base-pairs 36,168-37,605 (the DNA sequence of which is disclosed SEQ ID NO:206); insert 1/H6 spans base-pairs 36,055-37,354 (the DNA sequence of which is

disclosed in SEQ ID NO:207); and insert 3/C1 spans base-pairs 36,473-37,727 (the DNA sequence of which is disclosed in SEQ ID NO:208).

Each of these inserts encodes a fragment of the C-terminus of UL19 (also referred to as VP5 or ICP). The full length DNA and amino acid sequences for
5 UL19 are disclosed in SEQ ID NOs:210 and 212, respectively. The sequence shared by all three library inserts spans the corresponding HSV-2 (HG52 strain) genome at base-pairs 36,473-37,354 (the DNA and amino acid sequence of which is described in SEQ ID NOs:209 and 211, respectively).

The UL19 ORF spans the genomic sequence of HG52 at base-pairs
10 36,448-40,572, and encodes the major capsid protein of HSV-2. There are three nucleotide differences between the shared region encoded by the library inserts (derived from HSV-2 strain of 333) and the corresponding UL19 sequence from HG52: A36710G, G37248C, and C37317T (333/position/HG52). The first two nucleotide substitutions result in amino acid substitutions, the third does not. The substitutions are
15 N to S and G to A (HG52 to 333), respectively.

EXAMPLE 20

HSV-2 US8 GENE SEQUENCE FROM CLINICALLY ISOLATED VIRAL DNA

The DNA sequence corresponding to the US8 gene of HSV-2 was cloned from the laboratory HG52 viral strain and sequenced, the DNA and its corresponding
20 amino acid sequences are disclosed in SEQ ID NOs:118 and 120, respectively. The details of these experiments are described in Example 12. SEQ ID NO:118 was then compared to the HSV-2 HG52 strain genomic sequence contained in GenBank (accession number Z86099), the DNA and amino acid sequences of which are disclosed in SEQ ID NOs:119 and 121, respectively. This comparison revealed that SEQ ID
25 NO:118 contained an extra base pair at position 542 that resulted in a frameshift. The presence of this extra base pair was also confirmed in a second laboratory strain of HSV-2, 333. There was one additional base pair (bp 156) upstream of the first stop codon in SEQ ID NO:118 that differed from the GenBank US8 sequence (SEQ ID NO:119). No change in the US8 amino acid sequence would result from the change in
30 the nucleotide sequence at base pair 156. The full-length US8 gene was sequenced from

two clinical isolates; donors RW1874 and HV5101. In order to derive these sequences, PCR products from each clinical viral DNA isolate were cloned and ligated into a plasmid vector for sequencing. The vector contained an upstream fusion partner (Ra12/thrombin site) to enhance potential protein expression and provide a cleavage site, thus the clinical US8 gene sequences are lacking a starting ATG (Met) codon sequence so as to eliminate redundancy. The plasmid DNA containing each of these clones was sequenced using DNA primers specific for both vector and predicted internal sequence.

The full length US8 DNA and amino acid sequences for RW1874 are disclosed in SEQ ID NOs:213 and 215; and the full length US8 DNA and amino acid sequences for HV5101 are disclosed in SEQ ID NOs:214 and 216, respectively. These sequences were then compared with the published laboratory strain of HSV-2, HG52, the results of which are described in Table 5.

15

Table 5

nt position corresponding to HG52 sequence	RW1874/nt substitutions	RW1874/aa substitutions	HV5101/nt substitutions	HV5101/aa substitutions
1	ATG site removed	NA	ATG site removed	NA
129	9bp insert	Gly/Pro/Glu insert	9bp insert	Gly/Pro/Glu insert
551	Bp insert of G, resulting in frameshift	SerGluArgThr ProValSerVal ProProProThr	Bp insert of G, resulting in frameshift	SerGluArgThr ProValSerVal ProProAlaThr
584	bp deletion of C, results in a second frameshift that	-	bp deletion of C, results in a second frameshift that	-

	restores original reading frame		restores original reading frame	
681	T to C	Leu to Pro	T to C	Leu to Pro
814	T to G	Asp to Glu	T to G	Asp to Glu
1083	-	-	T to C	Val to Ala
1156	A to G	Glu to Glu	-	-
1221	A to C	His to Pro	-	-
1262	G to A	Val to Met	-	-

The differences observed in these clinical isolates provides valuable information on areas of sequences that are both highly conserved or demonstrate variability. This sequence information provides valuable information that can be exploited in the development of therapeutic or diagnostic antibodies for the treatment and prevention of HSV/HSV-2 infection. Variability in sequence of HSV-2 genes is largely unknown. This sequence information also provides valuable information for selection of an antigen or antigens for use in a vaccine that is to identify HSV-2 sequences that are most clinically representative and relevant.

10

EXAMPLE 21

IDENTIFICATION OF HSV-2 ANTIGENS USING CD4⁺ T CELL CLONING

This Example describes the further characterization of CD4⁺ T cell clones generated from donor HH, who is an HSV-2 exposed, but seronegative donor.

15 The generation of these T cell clones is described in detail in Example 6.

Clone HH6 was found to recognize UL21 presented by the HSV-2-III library, pool 3H11, in addition to pool D6 that was derived from the HSV2-II library. The insert DNA sequence and corresponding protein sequence are disclosed in SEQ ID NOs:217 and 227, respectively. These sequences were derived from the 333 strain of HSV2. The DNA sequence spans base pairs 42,908-44,296 of the homologous region in the HG52 viral genome. Full length DNA and protein sequences of UL21 derived

20

from the HG52 strain of HSV-2 are disclosed in SEQ ID NOs:218 and 228. Using 15-mer peptides overlapping by 10 amino acids, T cells from clone HH6 were tested for their ability to react when stimulated with the HSV-2 peptides in an ELISPOT assay. Results demonstrated that the reactive T cell epitope was located within the UL21 gene
5 between amino acids 281-300. The amino acids corresponding to this region of the sequence are: PLRELWWVFYAGDRALEEPH (SEQ ID NO:229).

Clone HH 20 was found to recognize a fragment of HSV-2 ORF, UL29. This clone was found to contain two UL29-encoding inserts, both of which were derived from the HG52 strain. The full length DNA and protein sequences of UL29
10 derived from the HG52 strain of HSV-2 are disclosed in SEQ ID NOs:221 and 232, respectively. The first insert, clone 1/C12_E1 (SEQ ID NO:219) spans base pairs 61,539-62,299 of the HG52 viral genome and encodes UL29 amino acids 48-303 (SEQ ID NO:230). The second insert, 2/E9_D11 (SEQ ID NO:220), spans base pairs 61,538-62360 of the HG52 viral genome and encodes UL29, amino acids 30-303 (SEQ ID NO:
15 231). The sequence of this insert differs from HG52 as follows: R121P and S126A (333/location/HG52).

Clone HH 22 recognized a fragment of the HSV-2 ORF of UL46 and was found to contain two inserts. The full-length DNA and amino acid sequences of UL46, derived from the HG52 strain of HSV-2, are disclosed in SEQ ID NOs:224 and
20 235, respectively. The first insert, F7_A1, derived from the HG52 strain of HSV-2, was found to span base pairs 99,253 to 100,014 of the homologous region in the HG52 viral genome and encodes UL46/amino acids 529-722. The DNA and amino acid sequence encoded by insert F7_A1 are disclosed in SEQ ID NOs:222 and 233, respectively. The amino acid sequence of the F7_A1 insert differs from HG52 as follows: -590A, S613G,
25 L643P, Q637R, D638L, P644L, P672R, G673R (333/location/HG52). The second insert, 4/E8_C8, derived from the HG52 strain of HSV-2 spans 99,232 to 100,262 of the HG52 viral genome and encodes UL46/amino acids 446-722. The DNA sequence of the insert and the amino acid sequence it encodes are disclosed in SEQ ID NOs:223 and 234, respectively. The location of a T cell epitope was mapped by screening 15-mer
30 peptides overlapping by 10 amino acid in an ELISPOT assay and is located within UL46/621-649 (EEIPWVRVYENICLRRQDA: SEQ ID NO:236).

Clone HH24 recognized a fragment of HSV-2 ORF UL47, the full-length DNA and amino acid sequences from the HG52 strain of HSV-2 are disclosed in SEQ ID NOs:226 and 238, respectively. The DNA sequence of the insert G6_H11 (SEQ ID NO:225 which encodes SEQ ID NO:237) spans base pairs 101,622 to 103,386 of the homologous region of the HG52 viral genome. The location of the T cell epitope was mapped by screening 15-mer peptides overlapping by 10 amino acids in an ELISPOT assay and is located with UL47/amino acids 137-155 (LGRVGGSRVVPSPFLDEL: SEQ ID NO:239).

EXAMPLE 22

10 IDENTIFICATION OF HSV-2 ANTIGENS USING CD4⁺ T CELL

CLONING

CD4⁺ T cells were generated from an HSV-2 seropositive donor, TM, who has suffered from infrequent genital lesion recurrences and sheds virus infrequently (as assessed by both virus cultures and PCR on daily swabs). TM clones were derived by stimulation of PBMCs with UV-inactivated HSV-2, strain 333, for 14 days, followed by limiting dilution cloning on anti-CD3 mAb. Clones were subsequently tested for reactivity with UV-inactivated HSV-2 and a panel of recombinant HSV-2 proteins.

Clones TM13 and TM58 recognize the same fragment of HSV-2 ORF UL54, also known as ICP27. The DNA and amino acid sequences corresponding to the insert 3/F5_G1 are disclosed in SEQ ID NOs:240 and 241, respectively. The corresponding HG52 DNA sequence spans base pairs 115,061 to 115,785 of the viral genome and encodes a fragment of UL54. The corresponding HG52 sequence for full-length UL54 is disclosed in SEQ ID NO:242. The actual amino acid sequence encoded by 3/F5_G1 (SEQ ID NO:241) corresponds to HG52 amino acids 159-399. The corresponding HG52 amino acid sequence for full-length UL54 is disclosed in SEQ ID NO:243. The 3/F5_G1 insert sequence differs from the HG52 UL54 sequence by a single amino acid, N169K (333/location/HG52).

Clone TM39 recognized an insert comprised of two genomic fragments of HSV-2. The DNA sequence for the insert 3/H11_C3 is disclosed in SEQ ID NO:244. The corresponding HG52 DNA sequences spanned include base pairs 43,717 to 44,086

and 70,294 to 71,846 of the viral genome. The first fragment, base pairs 70,294 to 71,846, encodes portions of HG52 UL21 and UL22. Base pairs 43,717 to 44,086, encodes portions of UL36. The corresponding full length DNA sequences for UL21, UL22, and UL36 are disclosed in SEQ ID NOs:245-247, respectively. The
5 corresponding full length HG52 amino acid sequences for UL21, UL22, and UL36 are disclosed in SEQ ID NOs:248-250, respectively.

Clone TM51 recognized a fragment of HSV-2 ORF US4. The DNA sequence of the insert derived from TM51, F7_A8, is disclosed in SEQ ID NO:251. The corresponding HG52 DNA sequence spans base pairs 139,505 to 140,104 of the
10 viral genome and encodes a fragment of US4. The corresponding HG52 DNA sequence is disclosed in SEQ ID NOs:252. The amino acid sequence encoded by F7_A8 is disclosed in SEQ ID NO:253, and corresponds to HG52 US4 amino acids 544-699. The corresponding hg52 amino acid sequence for the full length US4 sequence is disclosed in SEQ ID NO:254. The F7_A8 inset and HG52 amino acid sequence differ by a single
15 amino acid, D56E (333/location/HG52).

EXAMPLE 23

IDENTIFICATION OF CD4 AND CD8 T CELL EPITOPES FOR ICP0 AND UL47

Fifteen-mer peptides, overlapping by 11 amino acid residues, spanning
20 the ICP0 protein of HSV-2 were synthesized and arrayed into pools. BALB/c mice were immunized 3 times, at 3 week intervals with 100 ug of ICP0 plasmid DNA, with spleens harvested 2-3 weeks after the last immunization. CD8 T cells were purified from immune spleens by positive selection using MACS beads and columns. The cells from which CD8 T cells were purified were also saved and designated "CD8-depleted"
25 immune spleen cells.

Detection of CD8 T cell response was done using purified CD8 T cells (1.5×10^5 /well) co-cultured with naive BALB/c spleen APC (5×10^5 /well) to which the arrayed pools of ICP0 peptides (final concentration was 0.5 ug/ml of each peptide in pool) were added for a 24-hour murine IFN-g ELISPOT assay, medium alone served as
30 the negative control and Con A (5 ug/ml) served as the positive control.

Detection of CD4 T cell responses was carried out using CD8-depleted immune spleen cells (1×10^6 /well) co-cultured with the arrayed pools of 15-mer peptides (final concentration was 0.05 ug/ml of each peptide in pool) in a 24-hour murine IFN-g ELISPOT assay.

5 ELISPOT assays were developed and the results of the assays were determined both quantitatively and qualitatively. Two pools were found to be positive and the peptides from these pools #43 amino acid residues 211-225 of ICP0 (IWTGNPRTAPRSLSL SEQ ID NO: 255) and #44 amino acid residues 216-230 of ICP0 (PRTAPRSLSLGGHTV SEQ ID NO:256) shared 10 amino acid residues in
10 common. To identify the minimal ICP0 peptide recognized by CD8 T cells, we tested purified immune CD8 T cells against a series of synthetic ICP0 peptides (each 1 ug/ml) using the ELISPOT method described above. Amino acid residues 217-225 of ICP0 (RTAPRSLSL SEQ ID NO: 257) was determined to be the minimal epitope recognized by CD8 T cells

15 CD4 T cells recognized three peptides, #79 amino acid residues 390-404 of ICP0 (AQVSSGPGGGGLPQS SEQ ID NO:260), #82 amino acid residues 405-419 of ICP0 (SGRAARPRAAVAPRV SEQ ID NO: 259), and #89 440-454 of ICP0 PAPPAVPVDAHRAPR (SEQ ID NO: 258).

Fifteen-mer peptides, overlapping by 11 amino acid residues, spanning
20 the UL47 antigen of HSV-2 were synthesized and arrayed into pools and assayed as described above. Five peptides were identified from the positive pools, two of these #98 amino acid residues 389-403 of UL47 (YAGRMTYIATGALLA SEQ ID NO: 261) and #99 amino acid residues 393-407 of UL47 (MTYIATGALLARFNP SEQ ID NO:262), share a 10 amino acid residue overlap, amino acid residues 394-401 of UL47
25 (TYIATGALL SEQ ID NO: 263) was determined to be the minimal epitope recognized by CD8 T cells. Another peptide identified from a positive pool was #112 amino acid residues 445-459 of UL47 (ARLHPHSAHPAFADV SEQ ID NO:264). Amino acid residues 448-456 of of peptide #112 (HPHSAHPAF SEQ ID NO: 265) is considered the putative CD8 epitope.

30 CD4 T cells recognized peptide #1 amino acid residues 1-15 of UL47 (MSVRGHAVRRRRRAST SEQ ID NO:267), #4 amino acid residues 13-27 of UL47

(ASTRSHAPSAHRADS SEQ ID NO: 266), #99 amino acid residues 393-407 of UL 47 (MTYIATGALLARFNP SEQ ID NO: 262) and #112 amino acid residues 445-459 of UL 47 (ARLHPHSAHPAFADV SEQ ID NO: 264).

5 Those skilled in the art will appreciate that the conceptions and specific embodiments disclosed in the foregoing description may be readily utilized as a basis for modifying or designing other embodiments for carrying out the same purposes of the present invention. Those skilled in the art will also appreciate that such equivalent embodiments do not depart from the spirit and scope of the invention as set forth in the appended claims.

10 From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is Claimed:

1. An isolated polypeptide comprising at least an immunogenic portion of an HSV antigen, wherein said antigen comprises an amino acid sequence set forth in any one of SEQ ID NO: 161, 74-75, 90-97, 120-121, 122-140, 142-143, 153-160, 162-178, 181, 195-205, 215-216, 227-239, 241, 243, 248-250, 253-254 and 255-267.

2. An isolated polynucleotide encoding a polypeptide of claim 1.

3. An isolated polynucleotide of claim 2, wherein said polynucleotide comprises a sequence set forth in any one of SEQ ID NO: 152, 65-73, 76-89, 98-117, 118-119, 141, 144-151, 179-180 182-183, 184-194, 206-210, 213-214, 217-226, 240, 242, 244-247, and 251-252.

4. An isolated polypeptide comprising at least an immunogenic portion of the HSV UL19 antigen, wherein said antigen comprises an amino acid sequence set forth in SEQ ID NO: 212.

5. A fusion protein comprising a polypeptide according to claim 1 and a fusion partner.

6. A fusion protein according to claim 5, wherein the fusion partner comprises an expression enhancer that increases expression of the fusion protein in a host cell transfected with a polynucleotide encoding the fusion protein.

7. A fusion protein according to claim 5, wherein the fusion partner comprises a T helper epitope that is not present within the polypeptide of claim 1.

8. A fusion protein according to claim 5, wherein the fusion partner comprises an affinity tag.

9. An isolated polynucleotide encoding a fusion protein according to claim 5.

10. An isolated monoclonal or polyclonal antibody, or antigen-binding fragment thereof, that specifically binds to a polypeptide of claim 1.

11. A pharmaceutical composition comprising a polypeptide according to claim 1 or a polynucleotide encoding said polypeptide, and a physiologically acceptable carrier.

12. A pharmaceutical composition comprising a polypeptide according to claim 1, or a polynucleotide encoding said polypeptide, and an immunostimulant.

13. The pharmaceutical composition of claim 12, wherein the immunostimulant is selected from the group consisting of a monophosphoryl lipid A, aminoalkyl glucosaminide phosphate, saponin, or a combination thereof.

14. A method for stimulating an immune response in a patient, comprising administering to a patient a pharmaceutical composition according to any one of claims 11-13.

15. A method for detecting HSV infection in a patient, comprising:

- (a) obtaining a biological sample from the patient;
- (b) contacting the sample with a polypeptide according to claim 1;

and

- (c) detecting the presence of antibodies that bind to the polypeptide.

16. The method according to claim 15, wherein the biological sample is selected from the group consisting of whole blood, serum, plasma, saliva, cerebrospinal fluid and urine.

17. A method for detecting HSV infection in a biological sample, comprising:

(a) contacting the biological sample with a binding agent which is capable of binding to a polypeptide according to claim 1; and

(b) detecting in the sample a polypeptide that binds to the binding agent, thereby detecting HSV infection in the biological sample.

18. The method of claim 17, wherein the binding agent is a monoclonal antibody.

19. The method of claim 17, wherein the binding agent is a polyclonal antibody.

20. The method of claim 17 wherein the biological sample is selected from the group consisting of whole blood, sputum, serum, plasma, saliva, cerebrospinal fluid and urine.

21. A diagnostic kit comprising a component selected from the group consisting of:

(a) a polypeptide according to claim 1;

(b) a fusion protein according to claim 5;

(c) at least one antibody, or antigen-binding fragment thereof, according to claim 10; and

(d) a detection reagent.

22. The kit according to claim 21, wherein the polypeptide is immobilized on a solid support.

23. The kit according to claim 21, wherein the detection reagent comprises a reporter group conjugated to a binding agent.

24. The kit of claim 23, wherein the binding agent is selected from the group consisting of anti-immunoglobulins, Protein G, Protein A and lectins.

25. The kit of claim 23, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

SEQUENCE LISTING

<110> Corixa Corporation
Day, Craig H.
Hosken, Nancy A.
Parsons, Joseph M.

<120> COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
TREATMENT OF HERPES SIMPLEX VIRUS INFECTION

<130> 210121.53801PC

<140> PCT

<141> 2003-04-09

<160> 267

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 815

<212> DNA

<213> Herpes simplex virus

<400> 1

ccacgccgcc	gcacccagc	cgacgtggc	gccggttctg	gacagccagc	ccactgtggg	60
aacggacccc	ggctacccag	tccccctaga	actcacgccc	gagaacgcgg	aggcgggtggc	120
gcggtttctg	ggggacgccg	tgcacgcga	gccgcgctc	atgctggagt	acttctgtcg	180
gtgcgcccgc	gaggagagca	agcgcgtgcc	cccacgaacc	ttcggcagcg	ccccccgcct	240
cacggaggac	gactttgggc	tcctgaacta	cgcgtcgcct	gagatgcgac	gcctgtgcct	300
ggaccttccc	ccggtccccc	ccaacgcata	cacgccctat	catctgaggg	agtatgcgac	360
gcggtctggt	aacgggttca	aaccctgggt	gcggcggtcc	gcccgctgt	atcgcatcct	420
ggggattctg	gttcacctgc	gcatccgtac	ccgggaggcc	tcctttgagg	aatggatgcg	480
ctccaaggag	gtggacctgg	acttcgggct	gacggaaagg	cttcgcgaac	acgaggccca	540
gctaattgatc	ctggcccagg	ccctgaaccc	ctacgactgt	ctgatccaca	gcaccccgaa	600
cacgctcgtc	gagcgggggc	tgcagtcggc	gctgaagtac	gaagagtttt	acctcaagcg	660
cttcggcggg	cactacatgg	agtccgtctt	ccagatgtac	acccgcatcg	ccgggttcct	720
ggcgtgccgg	gcgaccgcg	gcatgcgcca	catcgccctg	gggcgacagg	ggtcgtggtg	780
ggaaatgttc	aagttctttt	tccaccgcct	ctacg			815

<210> 2

<211> 271

<212> PRT

<213> Herpes simplex virus

<400> 2

His	Ala	Ala	Ala	Pro	Gln	Ala	Asp	Val	Ala	Pro	Val	Leu	Asp	Ser	Gln
1				5				10					15		
Pro	Thr	Val	Gly	Thr	Asp	Pro	Gly	Tyr	Pro	Val	Pro	Leu	Glu	Leu	Thr
			20				25					30			
Pro	Glu	Asn	Ala	Glu	Ala	Val	Ala	Arg	Phe	Leu	Gly	Asp	Ala	Val	Asp
		35				40				45					
Arg	Glu	Pro	Ala	Leu	Met	Leu	Glu	Tyr	Phe	Cys	Arg	Cys	Ala	Arg	Glu
	50				55					60					
Glu	Ser	Lys	Arg	Val	Pro	Pro	Arg	Thr	Phe	Gly	Ser	Ala	Pro	Arg	Leu

```

65          70          75          80
Thr Glu Asp Asp Phe Gly Leu Leu Asn Tyr Ala Leu Ala Glu Met Arg
      85          90          95
Arg Leu Cys Leu Asp Leu Pro Pro Val Pro Pro Asn Ala Tyr Thr Pro
      100         105         110
Tyr His Leu Arg Glu Tyr Ala Thr Arg Leu Val Asn Gly Phe Lys Pro
      115         120         125
Leu Val Arg Arg Ser Ala Arg Leu Tyr Arg Ile Leu Gly Ile Leu Val
      130         135         140
His Leu Arg Ile Arg Thr Arg Glu Ala Ser Phe Glu Glu Trp Met Arg
      145         150         155         160
Ser Lys Glu Val Asp Leu Asp Phe Gly Leu Thr Glu Arg Leu Arg Glu
      165         170         175
His Glu Ala Gln Leu Met Ile Leu Ala Gln Ala Leu Asn Pro Tyr Asp
      180         185         190
Cys Leu Ile His Ser Thr Pro Asn Thr Leu Val Glu Arg Gly Leu Gln
      195         200         205
Ser Ala Leu Lys Tyr Glu Glu Phe Tyr Leu Lys Arg Phe Gly Gly His
      210         215         220
Tyr Met Glu Ser Val Phe Gln Met Tyr Thr Arg Ile Ala Gly Phe Leu
      225         230         235         240
Ala Cys Arg Ala Thr Arg Gly Met Arg His Ile Ala Leu Gly Arg Gln
      245         250         255
Gly Ser Trp Trp Glu Met Phe Lys Phe Phe Phe His Arg Leu Tyr
      260         265         270

```

<210> 3

<211> 1142

<212> PRT

<213> Herpes simplex virus

<400> 3

```

Met Ala Asn Arg Pro Ala Ala Ser Ala Leu Ala Gly Ala Arg Ser Pro
  1          5          10          15
Ser Glu Arg Gln Glu Pro Arg Glu Pro Glu Val Ala Pro Pro Gly Gly
      20         25         30
Asp His Val Phe Cys Arg Lys Val Ser Gly Val Met Val Leu Ser Ser
      35         40         45
Asp Pro Pro Gly Pro Ala Ala Tyr Arg Ile Ser Asp Ser Ser Phe Val
      50         55         60
Gln Cys Gly Ser Asn Cys Ser Met Ile Ile Asp Gly Asp Val Ala Arg
      65         70         75         80
Gly His Leu Arg Asp Leu Glu Gly Ala Thr Ser Thr Gly Ala Phe Val
      85         90         95
Ala Ile Ser Asn Val Ala Ala Gly Gly Asp Gly Arg Thr Ala Val Val
      100        105        110
Ala Leu Gly Gly Thr Ser Gly Pro Ser Ala Thr Thr Ser Val Gly Thr
      115        120        125
Gln Thr Ser Gly Glu Phe Leu His Gly Asn Pro Arg Thr Pro Glu Pro
      130        135        140
Gln Gly Pro Gln Ala Val Pro Pro Pro Pro Pro Pro Phe Pro Trp
      145        150        155        160
Gly His Glu Cys Cys Ala Arg Arg Asp Ala Arg Gly Gly Ala Glu Lys
      165        170        175
Asp Val Gly Ala Glu Ser Trp Ser Asp Gly Pro Ser Ser Asp Ser
      180        185        190
Glu Thr Glu Asp Ser Asp Ser Ser Asp Glu Asp Thr Gly Ser Glu Thr
      195        200        205

```

Leu Ser Arg Ser Ser Ser Ile Trp Ala Ala Gly Ala Thr Asp Asp Asp
 210 215 220
 Asp Ser Asp Ser Asp Ser Arg Ser Asp Asp Ser Val Gln Pro Asp Val
 225 230 235 240
 Val Val Arg Arg Arg Trp Ser Asp Gly Pro Ala Pro Val Ala Phe Pro
 245 250 255
 Lys Pro Arg Arg Pro Gly Asp Ser Pro Gly Asn Pro Gly Leu Gly Ala
 260 265 270
 Gly Thr Gly Pro Gly Ser Ala Thr Asp Pro Arg Ala Ser Ala Asp Ser
 275 280 285
 Asp Ser Ala Ala His Ala Ala Pro Gln Ala Asp Val Ala Pro Val
 290 295 300
 Leu Asp Ser Gln Pro Thr Val Gly Thr Asp Pro Gly Tyr Pro Val Pro
 305 310 315 320
 Leu Glu Leu Thr Pro Glu Asn Ala Glu Ala Val Ala Arg Phe Leu Gly
 325 330 335
 Asp Ala Val Asp Arg Glu Pro Ala Leu Met Leu Glu Tyr Phe Cys Arg
 340 345 350
 Cys Ala Arg Glu Glu Ser Lys Arg Val Pro Pro Arg Thr Phe Gly Ser
 355 360 365
 Ala Pro Arg Leu Thr Glu Asp Phe Gly Leu Leu Asn Tyr Ala Leu
 370 375 380
 Ala Glu Met Arg Arg Leu Cys Leu Asp Leu Pro Pro Val Pro Pro Asn
 385 390 395 400
 Ala Tyr Thr Pro Tyr His Leu Arg Glu Tyr Ala Thr Arg Leu Val Asn
 405 410 415
 Gly Phe Lys Pro Leu Val Arg Arg Ser Ala Arg Leu Tyr Arg Ile Leu
 420 425 430
 Gly Val Leu Val His Leu Arg Ile Arg Thr Arg Glu Ala Ser Phe Glu
 435 440 445
 Glu Trp Met Arg Ser Lys Glu Val Asp Leu Asp Phe Gly Leu Thr Glu
 450 455 460
 Arg Leu Arg Glu His Glu Ala Gln Leu Met Ile Leu Ala Gln Ala Leu
 465 470 475 480
 Asn Pro Tyr Asp Cys Leu Ile His Ser Thr Pro Asn Thr Leu Val Glu
 485 490 495
 Arg Gly Leu Gln Ser Ala Leu Lys Tyr Glu Glu Phe Tyr Leu Lys Arg
 500 505 510
 Phe Gly Gly His Tyr Met Glu Ser Val Phe Gln Met Tyr Thr Arg Ile
 515 520 525
 Ala Gly Phe Leu Ala Cys Arg Ala Thr Arg Gly Met Arg His Ile Ala
 530 535 540
 Leu Gly Arg Gln Gly Ser Trp Trp Glu Met Phe Lys Phe Phe Phe His
 545 550 555 560
 Arg Leu Tyr Asp His Gln Ile Val Pro Ser Thr Pro Ala Met Leu Asn
 565 570 575
 Leu Gly Thr Arg Asn Tyr Tyr Thr Ser Ser Cys Tyr Leu Val Asn Pro
 580 585 590
 Gln Ala Thr Thr Asn Gln Ala Thr Leu Arg Ala Ile Thr Gly Asn Val
 595 600 605
 Ser Ala Ile Leu Ala Arg Asn Gly Gly Ile Gly Leu Cys Met Gln Ala
 610 615 620
 Phe Asn Asp Ala Ser Pro Gly Thr Ala Ser Ile Met Pro Ala Leu Lys
 625 630 635 640
 Val Leu Asp Ser Leu Val Ala Ala His Asn Lys Gln Ser Thr Arg Pro
 645 650 655
 Thr Gly Ala Cys Val Tyr Leu Glu Pro Trp His Ser Asp Val Arg Ala
 660 665 670

Cys Thr Ser Cys Ala Leu
1140

<210> 4
<211> 208
<212> DNA
<213> Herpes simplex virus

<400> 4
gcgcgcgcgc cgcggtgccgc agaccacctc gcggcggtc cccgcgcgc tttcccggtg 60
ccctccacgc cgtggacgcc ccctcccaat tcgtcacctg gtcgcgcgtg cgctggctgc 120
ggggggcggt gggctctcggg gccgtcctgt gcgggattgc gttttacgtg acgtcaatcg 180
cccagagcgc ataaaggtcc ggcggcca 208

<210> 5
<211> 64
<212> PRT
<213> Herpes simplex virus

<400> 5
Gly Ala Ala Pro Ala Cys Arg Arg Pro Pro Arg Gly Gly Ser Pro Ala
1 5 10 15
Ala Phe Pro Val Ala Leu His Ala Val Asp Ala Pro Ser Gln Phe Val
20 25 30
Thr Trp Leu Ala Val Arg Trp Leu Arg Gly Ala Val Gly Leu Gly Ala
35 40 45
Val Leu Cys Gly Ile Ala Phe Tyr Val Thr Ser Ile Ala Arg Gly Ala
50 55 60

<210> 6
<211> 70
<212> PRT
<213> Herpes simplex virus

<400> 6
Arg Arg Ala Arg Val Pro Gln Thr Thr Ser Arg Arg Leu Pro Arg Gly
1 5 10 15
Leu Ser Arg Gly Pro Pro Arg Arg Gly Arg Pro Leu Pro Ile Arg His
20 25 30
Leu Ala Arg Arg Ala Leu Ala Ala Gly Gly Gly Gly Ser Arg Gly Arg
35 40 45
Pro Val Arg Asp Cys Val Leu Arg Asp Val Asn Arg Pro Arg Arg Ile
50 55 60
Lys Val Arg Arg Pro Ala
65 70

<210> 7
<211> 146
<212> PRT
<213> Herpes simplex virus

<400> 7
Met Asp Pro Ala Leu Arg Ser Tyr His Gln Arg Leu Arg Leu Tyr Thr
1 5 10 15
Pro Ile Ala Arg Gly Val Asn Leu Ala Ala Arg Ser Pro Pro Leu Val
20 25 30
Arg Glu Ala Arg Ala Val Val Thr Pro Arg Pro Pro Ile Arg Pro Ser
35 40 45

Ser Gly Lys Ala Ser Ser Asp Asp Ala Asp Val Gly Asp Glu Leu Ile
 50 55 60
 Ala Ile Ala Asp Ala Arg Gly Asp Pro Pro Glu Thr Leu Pro Pro Gly
 65 70 75 80
 Ala Gly Gly Ala Ala Pro Ala Cys Arg Arg Pro Pro Arg Gly Gly Ser
 85 90 95
 Pro Ala Ala Phe Pro Val Ala Leu His Ala Val Asp Ala Pro Ser Gln
 100 105 110
 Phe Val Thr Trp Leu Ala Val Arg Trp Leu Arg Gly Ala Val Gly Leu
 115 120 125
 Gly Ala Val Leu Cys Gly Ile Ala Phe Tyr Val Thr Ser Ile Ala Arg
 130 135 140
 Gly Ala
 145

<210> 8
 <211> 137
 <212> DNA
 <213> Herpes simplx virus

<400> 8
 cccacccgcc cccacacagg cggcgcggtgc ggagggcggc ccgtgcgtcc ccccggtccc 60
 cgcgggcggc ccgtggcgct cgggtgcccc ggtatggat tccgccccca acccggggtt 120
 tcgtggcctg cgtttcc 137

<210> 9
 <211> 430
 <212> DNA
 <213> Herpes simplex virus

<400> 9
 atggaccggg aggcacttcg ggccatcagc cgcgggtgca agcccccttc gaccctggca 60
 aaactggtga ccgggctggg attcgcgatc cacggagcgc tcatcccggg gtcggagggg 120
 tgtgtctttg atagcagcca ccgaactac cctcatcggt taatcgtaa ggcggggtgg 180
 tacgccagca cgaaccacga ggcgcggtct ctgagacgcc tgaaccaccc cgcgatccta 240
 cccctcctgg acctgcacgt cgtttctggg gtcacgtgtc tggctcctcc caagtatcac 300
 tgcgacctgt atacctatct gagcaagcgc ccgtctccgt tgggccacct acagataacc 360
 gcggtctccc ggcagctctt gagcgccatc gactacgtcc actgcgaagg catcatccac 420
 cgcgatatta 430

<210> 10
 <211> 22
 <212> PRT
 <213> Herpes simplex virus

<400> 10
 Trp Thr Gly Arg His Phe Gly Pro Ser Ala Ala Gly Ala Ser Pro Leu
 1 5 10 15
 Arg Pro Trp Gln Asn Trp
 20

<210> 11
 <211> 143
 <212> PRT
 <213> Herpes simplex virus

<400> 11
 Met Asp Arg Glu Ala Leu Arg Ala Ile Ser Arg Gly Cys Lys Pro Pro

```

1           5           10           15
Ser Thr Leu Ala Lys Leu Val Thr Gly Leu Gly Phe Ala Ile His Gly
20           25           30
Ala Leu Ile Pro Gly Ser Glu Gly Cys Val Phe Asp Ser Ser His Pro
35           40           45
Asn Tyr Pro His Arg Val Ile Val Lys Ala Gly Trp Tyr Ala Ser Thr
50           55           60
Asn His Glu Ala Arg Leu Leu Arg Arg Leu Asn His Pro Ala Ile Leu
65           70           75           80
Pro Leu Leu Asp Leu His Val Val Ser Gly Val Thr Cys Leu Val Leu
85           90           95
Pro Lys Tyr His Cys Asp Leu Tyr Thr Tyr Leu Ser Lys Arg Pro Ser
100          105          110
Pro Leu Gly His Leu Gln Ile Thr Ala Val Ser Arg Gln Leu Leu Ser
115          120          125
Ala Ile Asp Tyr Val His Cys Glu Gly Ile Ile His Arg Asp Ile
130          135          140

```

<210> 12

<211> 481

<212> PRT

<213> Herpes simplex virus

<400> 12

```

Met Ala Cys Arg Lys Phe Cys Gly Val Tyr Arg Arg Pro Asp Lys Arg
1           5           10           15
Gln Glu Ala Ser Val Pro Pro Glu Thr Asn Thr Ala Pro Ala Phe Pro
20           25           30
Ala Ser Thr Phe Tyr Thr Pro Ala Glu Asp Ala Tyr Leu Ala Pro Gly
35           40           45
Pro Pro Glu Thr Ile His Pro Ser Arg Pro Pro Ser Pro Gly Glu Ala
50           55           60
Ala Arg Leu Cys Gln Leu Gln Glu Ile Leu Ala Glu Met His Ser Asp
65           70           75           80
Glu Asp Tyr Pro Ile Val Asp Ala Ala Gly Ala Glu Glu Glu Asp Glu
85           90           95
Ala Asp Asp Asp Ala Pro Asp Asp Val Ala Tyr Pro Glu Asp Tyr Ala
100          105          110
Glu Gly Arg Phe Leu Ser Met Val Ser Ala Ala Pro Leu Pro Gly Ala
115          120          125
Ser Gly His Pro Pro Val Pro Gly Arg Ala Ala Pro Pro Asp Val Arg
130          135          140
Thr Cys Asp Thr Gly Lys Val Gly Ala Thr Gly Phe Thr Pro Glu Glu
145          150          155          160
Leu Asp Thr Met Asp Arg Glu Ala Leu Arg Ala Ile Ser Arg Gly Cys
165          170          175
Lys Pro Pro Ser Thr Leu Ala Lys Leu Val Thr Gly Leu Gly Phe Ala
180          185          190
Ile His Gly Ala Leu Ile Pro Gly Ser Glu Gly Cys Val Phe Asp Ser
195          200          205
Ser His Pro Asn Tyr Pro His Arg Val Ile Val Lys Ala Gly Trp Tyr
210          215          220
Ala Ser Thr Ser His Glu Ala Arg Leu Leu Arg Arg Leu Asn His Pro
225          230          235          240
Ala Ile Leu Pro Leu Leu Asp Leu His Val Val Ser Gly Val Thr Cys
245          250          255
Leu Val Leu Pro Lys Tyr His Cys Asp Leu Tyr Thr Tyr Leu Ser Lys
260          265          270

```

Arg Pro Ser Pro Leu Gly His Leu Gln Ile Thr Ala Val Ser Arg Gln
 275 280 285
 Leu Leu Ser Ala Ile Asp Tyr Val His Cys Lys Gly Ile Ile His Arg
 290 295 300
 Asp Ile Lys Thr Glu Asn Ile Phe Ile Asn Thr Pro Glu Asn Ile Cys
 305 310 315 320
 Leu Gly Asp Phe Gly Ala Ala Cys Phe Val Arg Gly Cys Arg Ser Ser
 325 330 335
 Pro Phe His Tyr Gly Ile Ala Gly Thr Ile Asp Thr Asn Ala Pro Glu
 340 345 350
 Val Leu Ala Gly Asp Pro Tyr Thr Gln Val Ile Asp Ile Trp Ser Ala
 355 360 365
 Gly Leu Val Ile Phe Glu Thr Ala Val His Thr Ala Ser Leu Phe Ser
 370 375 380
 Ala Pro Arg Asp Pro Glu Arg Arg Pro Cys Asp Asn Gln Ile Ala Arg
 385 390 395 400
 Ile Ile Arg Gln Ala Gln Val His Val Asp Glu Phe Pro Thr His Ala
 405 410 415
 Glu Ser Arg Leu Thr Ala His Tyr Arg Ser Arg Ala Ala Gly Asn Asn
 420 425 430
 Arg Pro Ala Trp Thr Arg Pro Ala Trp Thr Arg Tyr Tyr Lys Ile His
 435 440 445
 Thr Asp Val Glu Tyr Leu Ile Cys Lys Ala Leu Thr Phe Asp Ala Ala
 450 455 460
 Leu Arg Pro Ser Ala Ala Glu Leu Leu Arg Leu Pro Leu Phe His Pro
 465 470 475 480
 Lys

<210> 13

<211> 501

<212> DNA

<213> Herpes simplex virus

<400> 13

gggggcgcggt ctacgaggag atccctctggg ttcggttata cgaaaacatc tgccttcgcc 60
 ggcaagacgc cggcggggcg gccccgccgg gagacgcccc ggactccccg tacatcgagg 120
 cggaataatcc cctgtacgac tggggcggtgt ctgccctctt ctccctccg ggggccacac 180
 gcgccccgga cccgggacta agcctgtcgc ccatgccgc cgcgcccg accaacgcgc 240
 tggccaacga cggccccgaca aacgtcgccg ccctcagcgc cctgttgacg aagctcaaac 300
 gcggccgaca ccagagccat taaaaaaatg cgaccgcggg ccccaacgctc tcggtttccg 360
 gcccctttcc cgtatgtct gttttcaata aaaagtaaca aacagagaaa aaaaaacagc 420
 gagttccgca tggtttgtcg tacgcaatta gctgtttatt gttttttttt tggggggggg 480
 aagagaaaaa gaaaaaagga g 501

<210> 14

<211> 106

<212> PRT

<213> Herpes simplex virus

<400> 14

Gly Arg Val Tyr Glu Glu Ile Pro Trp Val Arg Val Tyr Glu Asn Ile
 1 5 10 15
 Cys Leu Arg Arg Gln Asp Ala Gly Gly Ala Ala Pro Pro Gly Asp Ala
 20 25 30
 Pro Asp Ser Pro Tyr Ile Glu Ala Glu Asn Pro Leu Tyr Asp Trp Gly
 35 40 45
 Gly Ser Ala Leu Phe Ser Pro Pro Gly Ala Thr Arg Ala Pro Asp Pro

50	55	60
Gly Leu Ser Leu Ser Pro Met Pro Ala Arg Pro Arg Thr Asn Ala Leu		
65	70	75
Ala Asn Asp Gly Pro Thr Asn Val Ala Ala Leu Ser Ala Leu Leu Thr		80
	85	90
Lys Leu Lys Arg Gly Arg His Gln Ser His		95
	100	105

<210> 15

<211> 722

<212> PRT

<213> Herpes simplex virus

<400> 15

Met Gln Arg Arg Ala Arg Gly Ala Ser Ser Leu Arg Leu Ala Arg Cys	
1	5
Leu Thr Pro Ala Asn Leu Ile Arg Gly Ala Asn Ala Gly Val Pro Glu	
	20
Arg Arg Ile Phe Ala Gly Cys Leu Leu Pro Thr Pro Glu Gly Leu Leu	
	35
Ser Ala Ala Val Gly Val Leu Arg Gln Arg Ala Asp Asp Leu Gln Pro	
	50
Ala Phe Leu Thr Gly Ala Asp Arg Ser Val Arg Leu Ala Ala Arg His	
65	70
His Asn Thr Val Pro Glu Ser Leu Ile Val Asp Gly Leu Ala Ser Asp	
	85
Pro His Tyr Asp Tyr Ile Arg His Tyr Ala Ser Ala Ala Lys Gln Ala	
	100
Leu Gly Glu Val Glu Leu Ser Gly Gly Gln Leu Ser Arg Ala Ile Leu	
	115
Ala Gln Tyr Trp Lys Tyr Leu Gln Thr Val Val Pro Ser Gly Leu Asp	
	130
Ile Pro Asp Asp Pro Ala Gly Asp Cys Asp Pro Ser Leu His Val Leu	
145	150
Leu Arg Pro Thr Leu Leu Pro Lys Leu Leu Val Arg Ala Pro Phe Lys	
	165
Ser Gly Ala Ala Ala Ala Lys Tyr Ala Ala Ala Val Ala Gly Leu Arg	
	180
Asp Ala Ala His Arg Leu Gln Gln Tyr Met Phe Phe Met Arg Pro Ala	
	195
Asp Pro Ser Arg Pro Ser Thr Asp Thr Ala Leu Arg Leu Ser Glu Leu	
	210
Leu Ala Tyr Val Ser Val Leu Tyr His Trp Ala Ser Trp Met Leu Trp	
225	230
Thr Ala Asp Lys Tyr Val Cys Arg Arg Leu Gly Pro Ala Asp Arg Arg	
	245
Phe Val Ala Leu Ser Gly Ser Leu Glu Ala Pro Ala Glu Thr Phe Ala	
	260
Arg His Leu Asp Arg Gly Pro Ser Gly Thr Thr Gly Ser Met Gln Cys	
	275
Met Ala Leu Arg Ala Ala Val Ser Asp Val Leu Gly His Leu Thr Arg	
	290
Leu Ala His Leu Trp Glu Thr Gly Lys Arg Ser Gly Gly Thr Tyr Gly	
305	310
Ile Val Asp Ala Ile Val Ser Thr Val Glu Val Leu Ser Ile Val His	
	325
His His Ala Gln Tyr Ile Ile Asn Ala Thr Leu Thr Gly Tyr Val Val	
	340

Trp Ala Ser Asp Ser Leu Asn Asn Glu Tyr Leu Thr Ala Ala Val Asp
 355 360 365
 Ser Gln Glu Arg Phe Cys Arg Thr Ala Ala Pro Leu Phe Pro Thr Met
 370 375 380
 Thr Ala Pro Ser Trp Ala Arg Met Glu Leu Ser Ile Lys Ser Trp Phe
 385 390 395 400
 Gly Ala Ala Leu Ala Pro Asp Leu Leu Arg Ser Gly Thr Pro Ser Pro
 405 410 415
 His Tyr Glu Ser Ile Leu Arg Leu Ala Ala Ser Gly Pro Pro Gly Gly
 420 425 430
 Arg Gly Ala Val Gly Gly Ser Cys Arg Asp Lys Ile Gln Arg Thr Arg
 435 440 445
 Arg Asp Asn Ala Pro Pro Pro Leu Pro Arg Ala Arg Pro His Ser Thr
 450 455 460
 Pro Ala Ala Pro Arg Arg Cys Arg Arg His Arg Glu Asp Leu Pro Glu
 465 470 475 480
 Pro Pro His Val Asp Ala Ala Asp Arg Gly Pro Glu Pro Cys Ala Gly
 485 490 495
 Arg Pro Ala Thr Tyr Tyr Thr His Met Ala Gly Ala Pro Pro Arg Leu
 500 505 510
 Pro Pro Arg Asn Pro Ala Pro Pro Glu Gln Arg Pro Ala Ala Ala Ala
 515 520 525
 Arg Pro Leu Ala Ala Gln Arg Glu Ala Ala Gly Val Tyr Asp Ala Val
 530 535 540
 Arg Thr Trp Gly Pro Asp Ala Glu Ala Glu Pro Asp Gln Met Glu Asn
 545 550 555 560
 Thr Tyr Leu Leu Pro Asp Asp Asp Ala Ala Met Pro Ala Gly Val Gly
 565 570 575
 Leu Gly Ala Thr Pro Ala Ala Asp Thr Thr Ala Ala Ala Trp Pro
 580 585 590
 Ala Glu Ser His Ala Pro Arg Ala Pro Ser Glu Asp Ala Asp Ser Ile
 595 600 605
 Tyr Glu Ser Val Gly Glu Asp Gly Gly Arg Val Tyr Glu Glu Ile Pro
 610 615 620
 Trp Val Arg Val Tyr Glu Asn Ile Cys Pro Arg Arg Arg Leu Ala Gly
 625 630 635 640
 Gly Ala Ala Leu Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala
 645 650 655
 Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro Arg
 660 665 670
 Arg Ala Thr Arg Ala Pro Asp Pro Gly Leu Ser Leu Ser Pro Met Pro
 675 680 685
 Ala Arg Pro Arg Thr Asn Ala Leu Ala Asn Asp Gly Pro Thr Asn Val
 690 695 700
 Ala Ala Leu Ser Ala Leu Leu Thr Lys Leu Lys Arg Gly Arg His Gln
 705 710 715 720
 Ser His

<210> 16

<211> 200

<212> DNA

<213> Herpes simplex virus

<400> 16

actgcaacgc aatcccatga aggccctgta tccgctcacc accaaggaac tcaagacttc	60
cgaccccggg ggcgtgggcg gggagggga ggaaggcgcg gagggggcg ggtttgacga	120
ggccaagttg gccgaggccc gagaaatgat ccgatatatg gctttggtgt cggccatgga	180

gcgcacggaa cacaaggcca

200

<210> 17
 <211> 66
 <212> PRT
 <213> Herpes simplex virus

<400> 17
 Leu Gln Arg Asn Pro Met Lys Ala Leu Tyr Pro Leu Thr Thr Lys Glu
 1 5 10 15
 Leu Lys Thr Ser Asp Pro Gly Gly Val Gly Gly Glu Gly Glu Glu Gly
 20 25 30
 Ala Glu Gly Gly Gly Phe Asp Glu Ala Lys Leu Ala Glu Ala Arg Glu
 35 40 45
 Met Ile Arg Tyr Met Ala Leu Val Ser Ala Met Glu Arg Thr Glu His
 50 55 60
 Lys Ala
 65

<210> 18
 <211> 904
 <212> PRT
 <213> Herpes simplex virus

<400> 18
 Met Arg Gly Gly Gly Leu Ile Cys Ala Leu Val Val Gly Ala Leu Val
 1 5 10 15
 Ala Ala Val Ala Ser Ala Ala Pro Ala Ala Pro Ala Ala Pro Arg Ala
 20 25 30
 Ser Gly Gly Val Ala Ala Thr Val Ala Ala Asn Gly Gly Pro Ala Ser
 35 40 45
 Arg Pro Pro Pro Val Pro Ser Pro Ala Thr Thr Lys Ala Arg Lys Arg
 50 55 60
 Lys Thr Lys Lys Pro Pro Lys Arg Pro Glu Ala Thr Pro Pro Pro Asp
 65 70 75 80
 Ala Asn Ala Thr Val Ala Ala Gly His Ala Thr Leu Arg Ala His Leu
 85 90 95
 Arg Glu Ile Lys Val Glu Asn Ala Asp Ala Gln Phe Tyr Val Cys Pro
 100 105 110
 Pro Pro Thr Gly Ala Thr Val Val Gln Phe Glu Gln Pro Arg Arg Cys
 115 120 125
 Pro Thr Arg Pro Glu Gly Gln Asn Tyr Thr Glu Gly Ile Ala Val Val
 130 135 140
 Phe Lys Glu Asn Ile Ala Pro Tyr Lys Phe Lys Ala Thr Met Tyr Tyr
 145 150 155 160
 Lys Asp Val Thr Val Ser Gln Val Trp Phe Gly His Arg Tyr Ser Gln
 165 170 175
 Phe Met Gly Ile Phe Glu Asp Arg Ala Pro Val Pro Phe Glu Glu Val
 180 185 190
 Ile Asp Lys Ile Asn Thr Lys Gly Val Cys Arg Ser Thr Ala Lys Tyr
 195 200 205
 Val Arg Asn Asn Met Glu Thr Thr Ala Phe His Arg Asp Asp His Glu
 210 215 220
 Thr Asp Met Glu Leu Lys Pro Ala Lys Val Ala Thr Arg Thr Ser Arg
 225 230 235 240
 Gly Trp His Thr Thr Asp Leu Lys Tyr Asn Pro Ser Arg Val Glu Ala
 245 250 255
 Phe His Arg Tyr Gly Thr Thr Val Asn Cys Ile Val Glu Glu Val Asp

260 265 270
 Ala Arg Ser Val Tyr Pro Tyr Asp Glu Phe Val Leu Ala Thr Gly Asp
 275 280 285
 Phe Val Tyr Met Ser Pro Phe Tyr Gly Tyr Arg Glu Gly Ser His Thr
 290 295 300
 Glu His Thr Ser Tyr Ala Ala Asp Arg Phe Lys Gln Val Asp Gly Phe
 305 310 315 320
 Tyr Ala Arg Asp Leu Thr Thr Lys Ala Arg Ala Thr Ser Pro Thr Thr
 325 330 335
 Arg Asn Leu Leu Thr Thr Pro Lys Phe Thr Val Ala Trp Asp Trp Val
 340 345 350
 Pro Lys Arg Pro Ala Val Cys Thr Met Thr Lys Trp Gln Glu Val Asp
 355 360 365
 Glu Met Leu Arg Ala Glu Tyr Gly Gly Ser Phe Arg Phe Ser Ser Asp
 370 375 380
 Ala Ile Ser Thr Thr Phe Thr Thr Asn Leu Thr Glu Tyr Ser Leu Ser
 385 390 395 400
 Arg Val Asp Leu Gly Asp Cys Ile Gly Arg Asp Ala Arg Glu Ala Ile
 405 410 415
 Asp Arg Met Phe Ala Arg Lys Tyr Asn Ala Thr His Ile Lys Val Gly
 420 425 430
 Gln Pro Gln Tyr Tyr Leu Ala Thr Gly Gly Phe Leu Ile Ala Tyr Gln
 435 440 445
 Pro Leu Leu Ser Asn Thr Leu Ala Glu Leu Tyr Val Arg Glu Tyr Met
 450 455 460
 Arg Glu Gln Asp Arg Lys Pro Arg Asn Ala Thr Pro Ala Pro Leu Arg
 465 470 475 480
 Glu Ala Pro Ser Ala Asn Ala Ser Val Glu Arg Ile Lys Thr Thr Ser
 485 490 495
 Ser Ile Glu Phe Ala Arg Leu Gln Phe Thr Tyr Asn His Ile Gln Arg
 500 505 510
 His Val Asn Asp Met Leu Gly Arg Ile Ala Val Ala Trp Cys Glu Leu
 515 520 525
 Gln Asn His Glu Leu Thr Leu Trp Asn Glu Ala Arg Lys Leu Asn Pro
 530 535 540
 Asn Ala Ile Ala Ser Ala Thr Val Gly Arg Arg Val Ser Ala Arg Met
 545 550 555 560
 Leu Gly Asp Val Met Ala Val Ser Thr Cys Val Pro Val Ala Pro Asp
 565 570 575
 Asn Val Ile Val Gln Asn Ser Met Arg Val Ser Ser Arg Pro Gly Thr
 580 585 590
 Cys Tyr Ser Arg Pro Leu Val Ser Phe Arg Tyr Glu Asp Gln Gly Pro
 595 600 605
 Leu Ile Glu Gly Gln Leu Gly Glu Asn Asn Glu Leu Arg Leu Thr Arg
 610 615 620
 Asp Ala Leu Glu Pro Cys Thr Val Gly His Arg Arg Tyr Phe Ile Phe
 625 630 635 640
 Gly Gly Gly Tyr Val Tyr Phe Glu Glu Tyr Ala Tyr Ser His Gln Leu
 645 650 655
 Ser Arg Ala Asp Val Thr Thr Val Ser Thr Phe Ile Asp Leu Asn Ile
 660 665 670
 Thr Met Leu Glu Asp His Glu Phe Val Pro Leu Glu Val Tyr Thr Arg
 675 680 685
 His Glu Ile Lys Asp Ser Gly Leu Leu Asp Tyr Thr Glu Val Gln Arg
 690 695 700
 Arg Asn Gln Leu His Asp Leu Arg Phe Ala Asp Ile Asp Thr Val Ile
 705 710 715 720
 Arg Ala Asp Ala Asn Ala Ala Met Phe Ala Gly Leu Cys Ala Phe Phe

725 730 735
 Glu Gly Met Gly Asp Leu Gly Arg Ala Val Gly Lys Val Val Met Gly
 740 745 750
 Val Val Gly Gly Val Val Ser Ala Val Ser Gly Val Ser Ser Phe Met
 755 760 765
 Ser Asn Pro Phe Gly Ala Leu Ala Val Gly Leu Leu Val Leu Ala Gly
 770 775 780
 Leu Val Ala Ala Phe Phe Ala Phe Arg Tyr Val Leu Gln Leu Gln Arg
 785 790 795 800
 Asn Pro Met Lys Ala Leu Tyr Pro Leu Thr Thr Lys Glu Leu Lys Thr
 805 810 815
 Ser Asp Pro Gly Gly Val Gly Gly Glu Gly Glu Glu Gly Ala Glu Gly
 820 825 830
 Gly Gly Phe Asp Glu Ala Lys Leu Ala Glu Ala Arg Glu Met Ile Arg
 835 840 845
 Tyr Met Ala Leu Val Ser Ala Met Glu Arg Thr Glu His Lys Ala Arg
 850 855 860
 Lys Lys Gly Thr Ser Ala Leu Leu Ser Ser Lys Val Thr Asn Met Val
 865 870 875 880
 Leu Arg Lys Arg Asn Lys Ala Arg Tyr Ser Pro Leu His Asn Glu Asp
 885 890 895
 Glu Ala Gly Asp Glu Asp Glu Leu
 900

<210> 19
 <211> 443
 <212> DNA
 <213> Herpes simplex virus

<400> 19
 ccctctccca cacggtcggt gccccccatc tctgtttcat catcggtcccgt gttgcgttgc 60
 gctttccggc cctcccgcac ccccgcggtc cgggtgtctcg cggcccggcg ccatgatcac 120
 ggattgtttc gaagcagaca tcgcgatccc ctcggtatc tcgcgccccg atgccgcggc 180
 gctgcagcgg tgcgagggtc gagtggtctt tctgccgacc atccgccgcc agctggcget 240
 cgcggacgtg gcgcacgaat cgttcgtctc cggaggagtt agtcccgcaca cgttgggggtt 300
 gttgctggcg taccgcaggc gttccccgc ggtaatcacg cgggtgctgc ccacgcgaat 360
 cgtcgcctgc cccgtggacc tggggctcac gcacgccggc accgtcaatc tccgcaacac 420
 ctccccgctc gacctctgca acg 443

<210> 20
 <211> 37
 <212> PRT
 <213> Herpes simplex virus

<400> 20
 Pro Leu Pro His Gly Arg Cys Pro Pro Ser Leu Phe His His Arg Pro
 1 5 10 15
 Gly Cys Val Ala Leu Ser Gly Pro Pro Ala Pro Pro Arg Ser Gly Val
 20 25 30
 Ser Arg Pro Gly Ala
 35

<210> 21
 <211> 147
 <212> PRT
 <213> Herpes simplex virus

<400> 21

Pro Leu Pro His Gly Arg Cys Pro Pro Ser Leu Phe His His Arg Pro
 1 5 10 15
 Gly Cys Val Ala Leu Ser Gly Pro Pro Ala Pro Pro Arg Ser Gly Val
 20 25 30
 Ser Arg Pro Gly Ala Met Ile Thr Asp Cys Phe Glu Ala Asp Ile Ala
 35 40 45
 Ile Pro Ser Gly Ile Ser Arg Pro Asp Ala Ala Ala Leu Gln Arg Cys
 50 55 60
 Glu Gly Arg Val Val Phe Leu Pro Thr Ile Arg Arg Gln Leu Ala Leu
 65 70 75 80
 Ala Asp Val Ala His Glu Ser Phe Val Ser Gly Gly Val Ser Pro Asp
 85 90 95
 Thr Leu Gly Leu Leu Leu Ala Tyr Arg Arg Arg Phe Pro Ala Val Ile
 100 105 110
 Thr Arg Val Leu Pro Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly
 115 120 125
 Leu Thr His Ala Gly Thr Val Asn Leu Arg Asn Thr Ser Pro Val Asp
 130 135 140
 Leu Cys Asn
 145

<210> 22
 <211> 110
 <212> PRT
 <213> Herpes simplex virus

<400> 22
 Met Ile Thr Asp Cys Phe Glu Ala Asp Ile Ala Ile Pro Ser Gly Ile
 1 5 10 15
 Ser Arg Pro Asp Ala Ala Ala Leu Gln Arg Cys Glu Gly Arg Val Val
 20 25 30
 Phe Leu Pro Thr Ile Arg Arg Gln Leu Ala Leu Ala Asp Val Ala His
 35 40 45
 Glu Ser Phe Val Ser Gly Gly Val Ser Pro Asp Thr Leu Gly Leu Leu
 50 55 60
 Leu Ala Tyr Arg Arg Arg Phe Pro Ala Val Ile Thr Arg Val Leu Pro
 65 70 75 80
 Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly Leu Thr His Ala Gly
 85 90 95
 Thr Val Asn Leu Arg Asn Thr Ser Pro Val Asp Leu Cys Asn
 100 105 110

<210> 23
 <211> 318
 <212> PRT
 <213> Herpes simplex virus

<400> 23
 Met Ile Thr Asp Cys Phe Glu Ala Asp Ile Ala Ile Pro Ser Gly Ile
 1 5 10 15
 Ser Arg Pro Asp Ala Ala Ala Leu Gln Arg Cys Glu Gly Arg Val Val
 20 25 30
 Phe Leu Pro Thr Ile Arg Arg Gln Leu Ala Leu Ala Asp Val Ala His
 35 40 45
 Glu Ser Phe Val Ser Gly Gly Val Ser Pro Asp Thr Leu Gly Leu Leu
 50 55 60
 Leu Ala Tyr Arg Arg Arg Phe Pro Ala Val Ile Thr Arg Val Leu Pro
 65 70 75 80

```

Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly Leu Thr His Ala Gly
      85          90          95
Thr Val Asn Leu Arg Asn Thr Ser Pro Val Asp Leu Cys Asn Gly Asp
      100          105          110
Pro Val Ser Leu Val Pro Pro Val Phe Glu Gly Gln Ala Thr Asp Val
      115          120          125
Arg Leu Glu Ser Leu Asp Leu Thr Leu Arg Phe Pro Val Pro Leu Pro
      130          135          140
Thr Pro Leu Ala Arg Glu Ile Val Ala Arg Leu Val Ala Arg Gly Ile
      145          150          155          160
Arg Asp Leu Asn Pro Asp Pro Arg Thr Pro Gly Glu Leu Pro Asp Leu
      165          170          175
Asn Val Leu Tyr Tyr Asn Gly Ala Arg Leu Ser Leu Val Ala Asp Val
      180          185          190
Gln Gln Leu Ala Ser Val Asn Thr Glu Leu Arg Ser Leu Val Leu Asn
      195          200          205
Met Val Tyr Ser Ile Thr Glu Gly Thr Thr Leu Ile Leu Thr Leu Ile
      210          215          220
Pro Arg Leu Leu Ala Leu Ser Ala Gln Asp Gly Tyr Val Asn Ala Leu
      225          230          235          240
Leu Gln Met Gln Ser Val Thr Arg Glu Ala Ala Gln Leu Ile His Pro
      245          250          255
Glu Ala Pro Met Leu Met Gln Asp Gly Glu Arg Arg Leu Pro Leu Tyr
      260          265          270
Glu Ala Leu Val Ala Trp Leu Ala His Ala Gly Gln Leu Gly Asp Ile
      275          280          285
Leu Ala Leu Ala Pro Ala Val Arg Val Cys Thr Phe Asp Gly Ala Ala
      290          295          300
Val Val Gln Ser Gly Asp Met Ala Pro Val Ile Arg Tyr Pro
      305          310          315

```

<210> 24
 <211> 502
 <212> DNA
 <213> Herpes simplex virus

```

<400> 24
actgttgtag ggggggaaaac acagttccgg gaaggcggtt attgcggaga gaggggggaa    60
agaaagagaa acaaaagaaa cggcaagaaa gactcaagac gtgcgcgtga tcggaaaaaa    120
ggccgggggg atcccgggtcg gggccgccag gtaaattggcc atgatgaccg cgaccatgag    180
gtcgtccgcg gcaccgttgc gttttccgga gtacatgcgg acgtcgggtg tgggagagac    240
ggttcgatg aggttggtga gctgctcgga cagatactcg accgggtcgg tctgcaggcg    300
caccgtcacg gagacgagct cctgggacgc catgacgccc ccggagtga actttttgat    360
aaagtattcg aaggcgggcg tcttctgttt gttgagcaga aagaaggggt acaataccgc    420
gccgccgggc ggctgcagc gatagaagag gagctcgggc cccgggccgt tggcccccgc    480
cgaggccagg atgcggtgca tc                                     502

```

<210> 25
 <211> 135
 <212> PRT
 <213> Herpes simplex virus

```

<400> 25
Met His Arg Ile Leu Ala Ser Ala Gly Ala Asn Gly Pro Gly Pro Glu
  1          5          10          15
Leu Leu Phe Tyr His Cys Glu Pro Pro Gly Gly Ala Val Leu Tyr Pro
      20          25          30
Phe Phe Leu Leu Asn Lys Gln Lys Thr Pro Ala Phe Glu Tyr Phe Ile

```

```

      35              40              45
Lys Lys Phe Asn Ser Gly Gly Val Met Ala Ser Gln Glu Leu Val Ser
  50              55              60
Val Thr Val Arg Leu Gln Thr Asp Pro Val Glu Tyr Leu Ser Glu Gln
  65              70              75              80
Leu Asn Asn Leu Ile Glu Thr Val Ser Pro Asn Thr Asp Val Arg Met
      85              90              95
Tyr Ser Gly Lys Arg Asn Gly Ala Ala Asp Asp Leu Met Val Ala Val
      100              105              110
Ile Met Ala Ile Tyr Leu Ala Ala Pro Thr Gly Ile Pro Pro Ala Phe
      115              120              125
Phe Pro Ile Thr Arg Thr Ser
      130              135

```

<210> 26
 <211> 734
 <212> PRT
 <213> Herpes simplex virus

```

<400> 26
Met Phe Gly Gln Gln Leu Ala Ser Asp Val Gln Gln Tyr Leu Glu Arg
  1              5              10              15
Leu Glu Lys Gln Arg Gln Gln Lys Val Gly Val Asp Glu Ala Ser Ala
      20              25              30
Gly Leu Thr Leu Gly Gly Asp Ala Leu Arg Val Pro Phe Leu Asp Phe
      35              40              45
Ala Thr Ala Thr Pro Lys Arg His Gln Thr Val Val Pro Gly Val Gly
      50              55              60
Thr Leu His Asp Cys Cys Glu His Ser Pro Leu Phe Ser Ala Val Ala
      65              70              75              80
Arg Arg Leu Leu Phe Asn Ser Leu Val Pro Ala Gln Leu Arg Gly Arg
      85              90              95
Asp Phe Gly Gly Asp His Thr Ala Lys Leu Glu Phe Leu Ala Pro Glu
      100              105              110
Leu Val Arg Ala Val Ala Arg Leu Arg Phe Arg Glu Cys Ala Pro Glu
      115              120              125
Asp Ala Val Pro Gln Arg Asn Ala Tyr Tyr Ser Val Leu Asn Thr Phe
      130              135              140
Gln Ala Leu His Arg Ser Glu Ala Phe Arg Gln Leu Val His Phe Val
      145              150              155              160
Arg Asp Phe Ala Gln Leu Leu Lys Thr Ser Phe Arg Ala Ser Ser Leu
      165              170              175
Ala Glu Thr Thr Gly Pro Pro Lys Lys Arg Ala Lys Val Asp Val Ala
      180              185              190
Thr His Gly Gln Thr Tyr Gly Thr Leu Glu Leu Phe Gln Lys Met Ile
      195              200              205
Leu Met His Ala Thr Tyr Phe Leu Ala Ala Val Leu Leu Gly Asp His
      210              215              220
Ala Glu Gln Val Asn Thr Phe Leu Arg Leu Val Phe Glu Ile Pro Leu
      225              230              235              240
Phe Ser Asp Thr Ala Val Arg His Phe Arg Gln Arg Ala Thr Val Phe
      245              250              255
Leu Val Pro Arg Arg His Gly Lys Thr Trp Phe Leu Val Pro Leu Ile
      260              265              270
Ala Leu Ser Leu Ala Ser Phe Arg Gly Ile Lys Ile Gly Tyr Thr Ala
      275              280              285
His Ile Arg Lys Ala Thr Glu Pro Val Phe Asp Glu Ile Asp Ala Cys
      290              295              300

```

Leu Arg Gly Trp Phe Gly Ser Ser Arg Val Asp His Val Lys Gly Glu
 305 310 315 320
 Thr Ile Ser Phe Ser Phe Pro Asp Gly Ser Arg Ser Thr Ile Val Phe
 325 330 335
 Ala Ser Ser His Asn Thr Asn Gly Ile Arg Gly Gln Asp Phe Asn Leu
 340 345 350
 Leu Phe Val Asp Glu Ala Asn Phe Ile Arg Pro Asp Ala Val Gln Thr
 355 360 365
 Ile Met Gly Phe Leu Asn Gln Ala Asn Cys Lys Ile Ile Phe Val Ser
 370 375 380
 Ser Thr Asn Thr Gly Lys Ala Ser Thr Ser Phe Leu Tyr Asn Leu Arg
 385 390 395 400
 Gly Ala Ala Asp Glu Leu Leu Asn Val Val Thr Tyr Ile Cys Asp Asp
 405 410 415
 His Met Pro Arg Val Val Thr His Thr Asn Ala Thr Ala Cys Ser Cys
 420 425 430
 Tyr Ile Leu Asn Lys Pro Val Phe Ile Thr Met Asp Gly Ala Val Arg
 435 440 445
 Arg Thr Ala Asp Leu Phe Leu Pro Asp Ser Phe Met Gln Glu Ile Ile
 450 455 460
 Gly Gly Gln Ala Arg Glu Thr Gly Asp Asp Arg Pro Val Leu Thr Lys
 465 470 475 480
 Ser Ala Gly Glu Arg Phe Leu Leu Tyr Arg Pro Ser Thr Thr Thr Asn
 485 490 495
 Ser Gly Leu Met Ala Pro Glu Leu Tyr Val Tyr Val Asp Pro Ala Phe
 500 505 510
 Thr Ala Asn Thr Arg Ala Ser Gly Thr Gly Ile Ala Val Val Gly Arg
 515 520 525
 Tyr Arg Asp Asp Phe Ile Ile Phe Ala Leu Glu His Phe Phe Leu Arg
 530 535 540
 Ala Leu Thr Gly Ser Ala Pro Ala Asp Ile Ala Arg Cys Val Val His
 545 550 555 560
 Ser Leu Ala Gln Val Leu Ala Leu His Pro Gly Ala Phe Arg Ser Val
 565 570 575
 Arg Val Ala Val Glu Gly Asn Ser Ser Gln Asp Ser Ala Val Ala Ile
 580 585 590
 Ala Thr His Val His Thr Glu Met His Arg Ile Leu Ala Ser Ala Gly
 595 600 605
 Ala Asn Gly Pro Gly Pro Glu Leu Leu Phe Tyr His Cys Glu Pro Pro
 610 615 620
 Gly Gly Ala Val Leu Tyr Pro Phe Phe Leu Leu Asn Lys Gln Lys Thr
 625 630 635 640
 Pro Ala Phe Glu Tyr Phe Ile Lys Lys Phe Asn Ser Gly Gly Val Met
 645 650 655
 Ala Ser Gln Glu Leu Val Ser Val Thr Val Arg Leu Gln Thr Asp Pro
 660 665 670
 Val Glu Tyr Leu Ser Glu Gln Leu Asn Asn Leu Ile Glu Thr Val Ser
 675 680 685
 Pro Asn Thr Asp Val Arg Met Tyr Ser Gly Lys Arg Asn Gly Ala Ala
 690 695 700
 Asp Asp Leu Met Val Ala Val Ile Met Ala Ile Tyr Leu Ala Ala Pro
 705 710 715 720
 Thr Gly Ile Pro Pro Ala Phe Phe Pro Ile Thr Arg Thr Ser
 725 730

<210> 27

<211> 15

<212> PRT

<213> HSV-2

<400> 27

Gly Arg Val Tyr Glu Glu Ile Pro Trp Val Arg Val Tyr Glu Asn
5 10 15

<210> 28

<211> 15

<212> PRT

<213> HSV-2

<400> 28

Tyr Glu Asn Ile Cys Leu Arg Arg Gln Asp Ala Gly Gly Ala Ala
5 10 15

<210> 29

<211> 15

<212> PRT

<213> HSV-2

<400> 29

Pro Asp Ser Pro Tyr Ile Glu Ala Glu Asn Pro Leu Tyr Asp Trp
5 10 15

<210> 30

<211> 15

<212> PRT

<213> HSV-2

<400> 30

Tyr Ile Glu Ala Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala
5 10 15

<210> 31

<211> 15

<212> PRT

<213> HSV-2

<400> 31

Thr Asn Ala Leu Ala Asn Asp Gly Pro Thr Asn Val Ala Ala Leu
5 10 15

<210> 32

<211> 15

<212> PRT

<213> HSV-2

<400> 32

Arg Val Leu Pro Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly
5 10 15

<210> 33

<211> 15
 <212> PRT
 <213> HSV-2

<400> 33
 Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly Leu Thr His Ala
 5 10 15

<210> 34
 <211> 661
 <212> DNA
 <213> HSV-2

<400> 34
 ctctctctcc gctctctcct cctctctctc egcctcctcc tctctctcct ccgctctctc 60
 ctctgcgggc ggggctggtg ggagcgctcg gtccgcgtcc ggcgctgggg agagacgaga 120
 aacctccctc ggcccccgcg ctgctgcgcc gggggggccg aggaagtgtg ccaggaagac 180
 gcgccaacgc gagggcgccc ccgagcccg ggcccgcgac ccggcgcccg gcctcacgcg 240
 ctacctgccc atcgcggggg tctcgagcgt cgtggccctg gcgccttacg tgaacaagac 300
 ggtcacgggg gactgcctgc ccgtcctgga catggagacg ggccacatag gggcctacgt 360
 ggtcctcggt gaccagacgg ggaacgtggc ggacctgctg cgggcccgcg ccccgcgtg 420
 gagccgcgcg acctgctcc ccgagcacgc gcgcaactgc gtgaggcccc ccgactacct 480
 gacgcccccc gcgtcggagt ggaacagcct ctggatgacc ccggtgggca acatgctctt 540
 tgaccagggc acctggtggt gcgcgctgga ctccacggc ctccggtcgc gccaccgctg 600
 gtctcgggag cagggcgcgc ccgcgcgggc cggcgacgcc cccgcggggc acggggagta 660
 g 661

<210> 35
 <211> 2481
 <212> DNA
 <213> HSV-2

<400> 35
 atggaacccc ggcccggaac gagctcccg ggcgacccc gccccgagcg gccgcccgcg 60
 cagacccccg gcacgcagcc cgcgcgcccg cagcctggg ggatgctcaa cgacatgcag 120
 tggctcgcca gcagcgactc ggaggaggag accgaggtgg gaatctctga cgacgacctt 180
 caccgcgact ccacctccga ggcgggcagc acggacacgg agatgttcga ggcgggcctg 240
 atggacgcgg ccacgcccc ggcccggccc ccggccgagc gccagggcag cccacgccc 300
 gccgacgcgc agggatcctg tgggggtggg ccctggtgtg aggaggaagc ggaagcggga 360
 gggggggggc acgtgtgtgc cgtgtgcacg gacgagatcg ccccgccctt gcgctgccag 420
 agttttccct gcctgcaccc ctctctgcatc ccgtgcatga agacctggat tccgttgccg 480
 aacacgtgtc ccctgtgcaa caccgcggtg gcgtacctga tagtgggcgt gaccgccagc 540
 gggctcgttca gcaccatccc gatagtgaac gacccccgga cccgcgtgga ggccgaggcg 600
 gccgtgcggg ccggcacggc cgtggacttt atctggacgg gcaacccgcg gacggccccg 660
 cgctccctgt cgctgggggg acacacggtc cgcgccctgt cgcaccccc cccgtggccc 720
 ggcacggacg acgaggacga tgacctggcc gacggtgtgg actacgtccc gcccgcccc 780
 cgaagagcgc cccggcgcgg gggcgggcgt gcgggggcga cccgcggaac ctcccagccc 840
 gccgcgaccc gaccggcgcc ccctggcgcc ccgcggagca gcagcagcgg cggcgccccg 900
 ttgcggggcg ggggtgggac tgggtctggg ggcggccctg ccgtcgcggc cgtcgtgccc 960
 agagtggcct ctcttcccc tgcggccggc ggggggcgcg cgcaggcgcg gcgggtgggc 1020
 gaagacgccg cggcggcgga gggcaggacg ccccccgga gacagccccg cgcggccccg 1080
 gageccccc tagtcatcag cgactctccc ccgcctctc cgcgcgcgcc cgcgggcccc 1140
 gggcgctct cctttgtctc ctctcctcc gcacaggtgt cctcgggccc cgggggggga 1200
 ggtctgccac agtcgtcggg gcgcgcgcgc cgcccccgcg cggcgtcgc cccgcgcgtc 1260
 cggagtccgc ccgcgcgcgc cgcgcgcgcc gtggtgtctg cgagcgcgga cgcggccggg 1320
 cccgcgcgc cgcgcgtgcc ggtggacgcg caccgcgcgc cccggtcgcg catgaccag 1380
 gctcagaccg acacccaagc acagagtctg ggccgggcag gcgcgaccga cgcgcgcggg 1440

```

tcgggaggggc cgggcgcgga gggaggagccc ggggtccccc ggggcaccaa cacccccggg 1500
gccgcccccc acgcccgcga gggggcggcg gcccgcctcc ggaagaggcg cgggtcggac 1560
tcgggccccg cggcctcgte ctccgcctct tectccgcg cccgcgctc gccctcgcc 1620
ccccaggggg tgggggcaa gaggcgcgcg ccgcgcggg ccccgactc ggactcgggc 1680
gaccgcggcc acgggcccgt cggcccggcg tccgcggcg ccgcgcccc gtcggcgctc 1740
ccgtcgtecc aggcgcgggt cgcgcgcgc tectctctc ccgcctctc ctctccgcc 1800
tctctctct cgcctctct ctctccgcc tectctctc ccgcctctc ctctccgcc 1860
tctctctct cgcctctct ctctcgggc ggggctggtg ggagcgctcg gtccgcgtcc 1920
ggcgctgggg agagacgaga aacctccctc ggcccccgcg ctgctgcgc gcgggggcgc 1980
aggaagtgtg ccaggaagac gcgccacgc gagggcggcg ccgagcccg ggcccgcgc 2040
ccggcgcccc gcctcacgcg ctacctgccc atcgcgggg tctcgagcgt cgtggccctg 2100
gcgccttacg tgaacaagac ggtcacgggg gactgcctgc ccgtcctgga catggagacg 2160
ggccacatag gggcctacgt ggtcctcggt gaccagacgg ggaacgtggc ggacctgctg 2220
cgggcccggg ccccgcggtg gagccgcgc acctgctcc ccgagcacgc gcgcaactgc 2280
gtgaggcccc ccgactaccc gacgcccccc gcgtcggagt ggaacagcct ctggatgacc 2340
ccggtgggca acatgctctt tgaccagggc acctgggtgg gcgcgctgga cttccacggc 2400
ctccggtcgc gccaccggtg gtctcgggag cagggcgcgc ccgcgcccgc cggcgacgcc 2460
cccgcgggcc acggggagta g

```

<210> 36
 <211> 1603
 <212> DNA
 <213> HSV-2

```

<400> 36
cggccggagg gctgtccgc atcgatatca cgagcccat gaagccctc ccgtatcgcg 60
cgcgcacgag cgcggcgctc caccgaacg ccagcccgcc cgtcgtccag acgcccacgg 120
gccacgtcga ggccgacggg gagaggatca cgtaccgacc cggagtccgt agcaggcccc 180
tggcgggccag ccaggtaacg gatgcgttgt gcagatgcgc gatgctcagg ttcgctcgte 240
gatgcctcgg tgtccccgcg gggcgccccg gggcgggcg gttgcgtcgg ccgtccgggt 300
gcctctcggt cggcccgctc tctccccgcg ggaacgtaag cccctcgcg tccgcgcggc 360
cgcgaatggt acccaggccc gggaccgcaa cagcgcgag gcgcccgggt tgtgcgacg 420
tcccttgagc tgggtcaact cggcgggggg acgggacgtg ggccccgct cggggagctc 480
gggcaggctc gcgttccgag gccggccgag cagataggctc tttgggatgt aaagcagctg 540
cccggggtcc cgaggaaact cggcgttgt gaccaacacg aaacaaaagc gtcggcgta 600
ccaccgaagc atgggcacgg atgcccgtgt caggttgagt tcgcccggg gcgccaagcg 660
tccgcgctgg ggtcgctgg cgtcgggggt tgttgggcaa ccacagacgc ccggtgtttt 720
gtcgcgccag tacgtcggg ccaacccag accgtgcaa aaccacgggt cgatttgctc 780
cgtccagtac gtgtcatggc ccccggaac gccaccagg acccccatca ccccccacag 840
accggggccc atggtcgctc gtcccggctg ccagtccgca gatggggggg ggtgtccgta 900
cccacggccc aaagaggctc cgcacctcg aggtatcgg aggcctttg ttgcccgaag 960
cgcgggcaaa aggatgggt ggggtgagg taaaagcaca aaggagtag cagaccgaaa 1020
acaaggacgg atcgccccg tccgttttct ggtgggtgc tgatacgtg ccagccctgg 1080
ccccgaaccc ccgcgttat ggacacacca cagacaaca atgcctttta ttctgttctt 1140
ttattgcgt catcgccggg aggccttccg ttcgggcttc cgtgtttgaa ctaaaactcc 1200
ccacctcgc gggcaaacgt gcgcgccagg tcgcgtatct cggcgatgga cccggcggtt 1260
gtgacgcggg ttgggatcat cccggcggtg aggcgcaaca gggcgctctc acaccgacg 1320
ggcgactgat cgtaatccag gacaaataga tgcacggaa ggaggcggtc ggccaagacg 1380
tccaagaccc aggcaaaaat gtggtacaag tccccgttg gggccagcag ctcggaacg 1440
cggaacaggg caaacagcgt gtcctcgat cggggcagag accccgcgc gtcctcgggg 1500
tcggggcgcg gggtcgcgc ggcgaccccc gtcagccggc ccagtcctc ccgccacctc 1560
ccgcgcgct gcaggtaacc caccgtgttg gcgagtagat cgt

```

<210> 37
 <211> 1131
 <212> DNA
 <213> HSV-2

<400> 37

```

atggcttctc acgccggcca acagcacgcg cctgcgttcg gtcaggctgc tcgtgcgagc 60
gggcctaccg acggccgcgc ggcgtcccgt cctagccatc gccagggggc ctccggagcc 120
cgccggggatc cggagctgcc cacgctgctg cgggtttata tagacggacc ccacgggggtg 180
gggaagacca ccacctccgc gcagctgatg gaggccctgg ggccgcgcga caatatcgtc 240
tacgtccccg agccgatgac ttactggcag gtgctggggg cctccgagac cctgacgaac 300
atctacaaca cgcagcacgc tctggaccgc ggcgagatat cggccgggga ggcggcggtg 360
gtaatgacca gcgccagat aacaatgagc acgccttatg cggcgacgga cgccgttttg 420
gctcctcata tcggggggga ggctgtgggc ccgcaagccc cgccccggc cctcaccctt 480
gttttcgacc ggcacctat cgctccctg ctgtgctacc cggccgcgcg gtacctcatg 540
ggaagcatga cccccaggc cgtgttggcg ttctggccc tcatgcccc gaccgcgcc 600
ggcacgaacc tggctctggg tgtccttcgc gaggccgaac acgccgaccg cctggccaga 660
cgccaacgcc cggcgagcgc gcttgacctg gccatgctgt ccgccattcg ccgtgtctac 720
gatctactcg ccaacacggt gcggtacctg cagcgcggcg ggaggtggcg ggaggactgg 780
ggccggctga cgggggtcgc cgcggcgacc ccgcgcccc .accccgagga cggcgcgggg 840
tctctgcccc gcctcgagga cacgtgttt gccctgttcc gcgttccga gctgctggcc 900
cccaacgggg acttgtacca catTTTTgccc tgggtcttgg acgtcttggc cgaccgcctc 960
cttccgatgc atctatttgt cctggattac gatcagtcgc ccgtcgggtg tcgagacgcc 1020
ctgttgccgc tcaccgccgc gatgatccca acccgcgctc caaccgccgc gtccatcgcc 1080
gagatacgcg acctggcgcg cacgtttgccc cgcgaggtgg ggggagttaa g 1131

```

<210> 38

<211> 2517

<212> DNA

<213> HSV-2

<400> 38

```

atgggccccg gtctgtgggt ggtgatgggg gtccctgggtg gcgttgccgg gggccatgac 60
acgtactgga cggagcaaat cgaccctggg tttttgcacg gtctgggggt ggcccgcacg 120
tactggcgcg acacaaacac cgggcgtctg tggttgcca acacccccga cgccagcgac 180
ccccagcgcg gacgcttggc gccccggggc gaactcaacc tgactacggc atccgtgccc 240
atgcttcggt ggtacgcga gcgcttttgt ttctgttgg tcaccacggc cgagtttctt 300
cgggaccccc ggcagctgct ttacatccca aagacctatc tgctcgccg gcctcggaac 360
gcgagcctgc ccgagctccc cgaggcgggg ccacgtccc gtccccccgc cgaggtgacc 420
cagctcaagg gactgtcgca caaccccggc gcctccgcgc tgttgcggtc ccgggcctgg 480
gtaacattcg cggccgcgcg ggaccgcgag gggcttacgt tcccgcgggg agacgacggg 540
gcgaccgaga ggcacccgga cggccgacgc aacgcgcgcg ccccgggggc gcccgcgggg 600
acaccgaggc atccgacgac gaacctgagc atcgcgcatc tgcaaacgc atccgtgacc 660
tggttgcccg ccaggggcct gctacggact ccgggtcgt acgtgtacct ctcccgtcg 720
gcctcgacgt ggcccggtgg cgtctggacg acgggcgggc tggcggttcg gtgcgacgcc 780
gcgctcgtgc gcgcgcgata cgggaagggc ttcatggggc tcgtgatata gatcggggac 840
agcctccgg cggagatcat agtggtgct gcggacaaga ccctcgctcg ggtcggaaat 900
ccgaccgacg aaaacgcccc cgcggtgct cccgggcctc cggccggccc caggtatcgc 960
gtctttgtcc tggggggccc gacgccgcg gacaacggct cggcgctgga cgccctccg 1020
cgggtggccg gctacccccga ggagagcacg aactacgccc agtatatgt cggggcctat 1080
gcggagtttt tgggggagga cccgggctcc ggcacggacg cgcgtccgtc cctgttctgg 1140
cgctcgcgg ggctgctcgc ctgctcgggg tttgcgttcg tcaacgcggc ccacgcccac 1200
gacgcgattc gcctctccga cctgctgggc tttttggccc actcgcgcgt gctggccggc 1260
ctggccgccc ggggagcagc gggctgcgcg gccgactcgg tgttctgaa cgtgtccgtg 1320
ttggaccgg cgcccgccct gcggctggag gcgcgcctcg ggcatctggt ggccgcgac 1380
ctcgagcgag agcagagcct ggtggcgcac gcgctgggt atcagctggc gttcgtgttg 1440
gacagccccg cggcctatgg cgcggtggcc ccgagcgcg cccgcctgat cgacgccctg 1500
tacgccaggt ttctcggcgg ccgcgcgcta accgccccga tggtcgcgc agcgtgttt 1560
tacgccagcg ccgtcctccg ggcgcgctt ctggcgggcg cgccctcggc cgagcagcgg 1620
gaacgcgccc ccggggcct cctcataaac acggccctgt gtacgtccga cgtcgcgcg 1680
gcgacccacg cccgattccg ggccgcgcta gccaggaccg accaccagaa aaacctctt 1740
tggctccgg accacttttc ccatgcgca gcttccctgc gcttcgatct cgccgagggc 1800
gggttcatcc tggacgcgct ggccatggcc acccgatccg acatcccggc ggacgtcatg 1860

```



```

gcacaacaga cccgcggcgt ggcctccgtt ctcacgcgct gggcgcaacta caacgccctg 1920
atccgcgcct tcgtcccga ggccaccac cagtgtacg gcccgctcga caacgcggag 1980
ccccggatcc tcgtgccat caccacaac gccagctacg tcgtcaccca caccoccttg 2040
ccccgcggga tcggatacaa gcttacgggc gttgacgtcc gccgcccgt gtttatcacc 2100
tatctcaccg ccacctgcga agggcacgcg cgggagattg agccgaagcg gctgggtgcgc 2160
accgaaaacc ggcgcgacct cggcctcggt gggggcgtgt ttctgcgcta caccocggcc 2220
ggggaggtca tgctgggtgt gctgggtggac acggatgcc aaccaacagca gctggcccag 2280
gggcgggtgg cgggcacccc gaacgtgttt tccagcgacg tgccgtccgt ggccctgttg 2340
ttgttcccca acggaactgt gattcatctg ctggcctttg acacgtgcc catcgccacc 2400
atcgcccccg ggtttctggc cgcgtccgcg ctgggggtcg ttatgattac cgcggccctg 2460
gcgggcatcc ttaggggtgt ccgaacgtgc gtcccathtt tgtggagacg cgaataa 2517

```

<210> 39
 <211> 376
 <212> PRT
 <213> HSV-2

<400> 39

Met	Ala	Ser	His	Ala	Gly	Gln	Gln	His	Ala	Pro	Ala	Phe	Gly	Gln	Ala	5	10	15
Ala	Arg	Ala	Ser	Gly	Pro	Thr	Asp	Gly	Arg	Ala	Ala	Ser	Arg	Pro	Ser	20	25	30
His	Arg	Gln	Gly	Ala	Ser	Gly	Ala	Arg	Gly	Asp	Pro	Glu	Leu	Pro	Thr	35	40	45
Leu	Leu	Arg	Val	Tyr	Ile	Asp	Gly	Pro	His	Gly	Val	Gly	Lys	Thr	Thr	50	55	60
Thr	Ser	Ala	Gln	Leu	Met	Glu	Ala	Leu	Gly	Pro	Arg	Asp	Asn	Ile	Val	65	70	75
Tyr	Val	Pro	Glu	Pro	Met	Thr	Tyr	Trp	Gln	Val	Leu	Gly	Ala	Ser	Glu	80	85	90
Thr	Leu	Thr	Asn	Ile	Tyr	Asn	Thr	Gln	His	Arg	Leu	Asp	Arg	Gly	Glu	95	100	105
Ile	Ser	Ala	Gly	Glu	Ala	Ala	Val	Val	Met	Thr	Ser	Ala	Gln	Ile	Thr	110	115	120
Met	Ser	Thr	Pro	Tyr	Ala	Ala	Thr	Asp	Ala	Val	Leu	Ala	Pro	His	Ile	125	130	135
Gly	Gly	Glu	Ala	Val	Gly	Pro	Gln	Ala	Pro	Pro	Pro	Ala	Leu	Thr	Leu	140	145	150
Val	Phe	Asp	Arg	His	Pro	Ile	Ala	Ser	Leu	Leu	Cys	Tyr	Pro	Ala	Ala	155	160	165
Arg	Tyr	Leu	Met	Gly	Ser	Met	Thr	Pro	Gln	Ala	Val	Leu	Ala	Phe	Val	170	175	180
Ala	Leu	Met	Pro	Pro	Thr	Ala	Pro	Gly	Thr	Asn	Leu	Val	Leu	Gly	Val	185	190	195
Leu	Pro	Glu	Ala	Glu	His	Ala	Asp	Arg	Leu	Ala	Arg	Arg	Gln	Arg	Pro	200	205	210
Gly	Glu	Arg	Leu	Asp	Leu	Ala	Met	Leu	Ser	Ala	Ile	Arg	Arg	Val	Tyr	215	220	225
Asp	Leu	Leu	Ala	Asn	Thr	Val	Arg	Tyr	Leu	Gln	Arg	Gly	Gly	Arg	Trp	230	235	240
Arg	Glu	Asp	Trp	Gly	Arg	Leu	Thr	Gly	Val	Ala	Ala	Ala	Thr	Pro	Arg	245	250	255
Pro	Asp	Pro	Glu	Asp	Gly	Ala	Gly	Ser	Leu	Pro	Arg	Ile	Glu	Asp	Thr	260	265	270
Leu	Phe	Ala	Leu	Phe	Arg	Val	Pro	Glu	Leu	Leu	Ala	Pro	Asn	Gly	Asp	275	280	285
Leu	Tyr	His	Ile	Phe	Ala	Trp	Val	Leu	Asp	Val	Leu	Ala	Asp	Arg	Leu	290	295	300
																305	310	315

<210>	40
<211>	136
<212>	PRT
<213>	HSV-2

```
<210> 41
<211> 284
<212> PRT
<213> HSV-2
```

<400> 41																
Met	Gly	Pro	Gly	Leu	Trp	Val	Val	Met	Gly	Val	Leu	Val	Gly	Val	Ala	
				5					10						15	
Gly	Gly	His	Asp	Thr	Tyr	Trp	Thr	Glu	Gln	Ile	Asp	Pro	Trp	Phe	Leu	
			20					25					30			
His	Gly	Leu	Gly	Leu	Ala	Arg	Thr	Tyr	Trp	Arg	Asp	Thr	Asn	Thr	Gly	
		35					40					45				
Arg	Leu	Trp	Leu	Pro	Asn	Thr	Pro	Asp	Ala	Ser	Asp	Pro	Gln	Arg	Gly	
	50					55					60					
Arg	Leu	Ala	Pro	Pro	Gly	Glu	Leu	Asn	Leu	Thr	Thr	Ala	Ser	Val	Pro	
	65				70					75					80	
Met	Leu	Arg	Trp	Tyr	Ala	Glu	Arg	Phe	Cys	Phe	Val	Leu	Val	Thr	Thr	
				85					90						95	
Ala	Glu	Phe	Pro	Arg	Asp	Pro	Gly	Gln	Leu	Leu	Tyr	Ile	Pro	Lys	Thr	
			100					105					110			
Tyr	Leu	Leu	Gly	Arg	Pro	Arg	Asn	Ala	Ser	Leu	Pro	Glu	Leu	Pro	Glu	
		115					120					125				
Ala	Gly	Pro	Thr	Ser	Arg	Pro	Pro	Ala	Glu	Val	Thr	Gln	Leu	Lys	Gly	

130		135		140	
Leu Ser His Asn Pro Gly Ala Ser Ala Leu Leu Arg Ser Arg Ala Trp					
145		150		155	160
Val Thr Phe Ala Ala Ala Pro Asp Arg Glu Gly Leu Thr Phe Pro Arg					
	165		170		175
Gly Asp Asp Gly Ala Thr Glu Arg His Pro Asp Gly Arg Arg Asn Ala					
	180		185		190
Pro Pro Pro Gly Pro Pro Ala Gly Thr Pro Arg His Pro Thr Thr Asn					
	195		200		205
Leu Ser Ile Ala His Leu His Asn Ala Ser Val Thr Trp Leu Ala Ala					
	210		215		220
Arg Gly Leu Leu Arg Thr Pro Gly Arg Tyr Val Tyr Leu Ser Pro Ser					
225		230		235	240
Ala Ser Thr Trp Pro Val Gly Val Trp Thr Thr Gly Gly Leu Ala Phe					
	245		250		255
Gly Cys Asp Ala Ala Leu Val Arg Ala Arg Tyr Gly Lys Gly Phe Met					
	260		265		270
Gly Leu Val Ile Ser Met Arg Asp Ser Pro Pro Ala					
	275		280		

<210> 42
 <211> 15
 <212> PRT
 <213> HSV-2

<400> 42
 Ser Leu Pro Arg Ile Glu Asp Thr Leu Phe Ala Leu Phe Arg Val
 5 10 15

<210> 43
 <211> 15
 <212> PRT
 <213> HSV-2

<400> 43
 Gly Ser Ile Ala Glu Ile Arg Asp Leu Ala Arg Thr Phe Ala Arg
 5 10 15

<210> 44
 <211> 16
 <212> PRT
 <213> HSV-2

<400> 44
 Glu Ile Arg Asp Leu Ala Arg Thr Phe Ala Arg Glu Val Gly Gly Val
 5 10 15

<210> 45
 <211> 838
 <212> PRT
 <213> HSV-2

<400> 45
 Met Gly Pro Gly Leu Trp Val Val Met Gly Val Leu Val Gly Val Ala
 5 10 15
 Gly Gly His Asp Thr Tyr Trp Thr Glu Gln Ile Asp Pro Trp Phe Leu


```

      485      490      495
Ile Asp Ala Leu Tyr Ala Glu Phe Leu Gly Gly Arg Ala Leu Thr Ala
      500      505      510
Pro Met Val Arg Arg Ala Leu Phe Tyr Ala Thr Ala Val Leu Arg Ala
      515      520      525
Pro Phe Leu Ala Gly Ala Pro Ser Ala Glu Gln Arg Glu Arg Ala Arg
      530      535      540
Arg Gly Leu Leu Ile Thr Thr Ala Leu Cys Thr Ser Asp Val Ala Ala
545      550      555      560
Ala Thr His Ala Asp Leu Arg Ala Ala Leu Ala Arg Thr Asp His Gln
      565      570      575
Lys Asn Leu Phe Trp Leu Pro Asp His Phe Ser Pro Cys Ala Ala Ser
      580      585      590
Leu Arg Phe Asp Leu Ala Glu Gly Gly Phe Ile Leu Asp Ala Leu Ala
      595      600      605
Met Ala Thr Arg Ser Asp Ile Pro Ala Asp Val Met Ala Gln Gln Thr
      610      615      620
Arg Gly Val Ala Ser Val Leu Thr Arg Trp Ala His Tyr Asn Ala Leu
625      630      635      640
Ile Arg Ala Phe Val Pro Glu Ala Thr His Gln Cys Ser Gly Pro Ser
      645      650      655
His Asn Ala Glu Pro Arg Ile Leu Val Pro Ile Thr His Asn Ala Ser
      660      665      670
Tyr Val Val Thr His Thr Pro Leu Pro Arg Gly Ile Gly Tyr Lys Leu
      675      680      685
Thr Gly Val Asp Val Arg Arg Pro Leu Phe Ile Thr Tyr Leu Thr Ala
      690      695      700
Thr Cys Glu Gly His Ala Arg Glu Ile Glu Pro Lys Arg Leu Val Arg
705      710      715      720
Thr Glu Asn Arg Arg Asp Leu Gly Leu Val Gly Ala Val Phe Leu Arg
      725      730      735
Tyr Thr Pro Ala Gly Glu Val Met Ser Val Leu Leu Val Asp Thr Asp
      740      745      750
Ala Thr Gln Gln Gln Leu Ala Gln Gly Pro Val Ala Gly Thr Pro Asn
      755      760      765
Val Phe Ser Ser Asp Val Pro Ser Val Ala Leu Leu Leu Phe Pro Asn
      770      775      780
Gly Thr Val Ile His Leu Leu Ala Phe Asp Thr Leu Pro Ile Ala Thr
785      790      795      800
Ile Ala Pro Gly Phe Leu Ala Ala Ser Ala Leu Gly Val Val Met Ile
      805      810      815
Thr Ala Ala Leu Ala Gly Ile Leu Arg Val Val Arg Thr Cys Val Pro
      820      825      830
Phe Leu Trp Arg Arg Glu
      835

```

<210> 46
 <211> 215
 <212> PRT
 <213> HSV-2

<400> 46
 Ser Ser Ser Ala Ser Ser Ser Ala Ser Ser Ser Ser Ala Ser Ser
 5 10 15
 Ser Ala Gly Gly Ala Gly Gly Ser Val Ala Ser Ala Ser Gly Ala Gly
 20 25 30
 Glu Arg Arg Glu Thr Ser Leu Gly Pro Arg Ala Ala Ala Pro Arg Gly
 35 40 45

Pro Arg Lys Cys Ala Arg Lys Thr Arg His Ala Glu Gly Gly Pro Glu
 50 55 60
 Pro Gly Ala Arg Asp Pro Ala Pro Gly Leu Thr Arg Tyr Leu Pro Ile
 65 70 75 80
 Ala Gly Val Ser Ser Val Val Ala Leu Ala Pro Tyr Val Asn Lys Thr
 85 90 95
 Val Thr Gly Asp Cys Leu Pro Val Leu Asp Met Glu Thr Gly His Ile
 100 105 110
 Gly Ala Tyr Val Val Leu Val Asp Gln Thr Gly Asn Val Ala Asp Leu
 115 120 125
 Leu Arg Ala Ala Ala Pro Ala Trp Ser Arg Arg Thr Leu Leu Pro Glu
 130 135 140
 His Ala Arg Asn Cys Val Arg Pro Pro Asp Tyr Pro Thr Pro Pro Ala
 145 150 155 160
 Ser Glu Trp Asn Ser Leu Trp Met Thr Pro Val Gly Asn Met Leu Phe
 165 170 175
 Asp Gln Gly Thr Leu Val Gly Ala Leu Asp Phe His Gly Leu Arg Ser
 180 185 190
 Arg His Pro Trp Ser Arg Glu Gln Gly Ala Pro Ala Pro Ala Gly Asp
 195 200 205
 Ala Pro Ala Gly His Gly Glu
 210 215

<210> 47
 <211> 826
 <212> PRT
 <213> HSV-2

<400> 47
 Met Glu Pro Arg Pro Gly Thr Ser Ser Arg Ala Asp Pro Gly Pro Glu
 5 10 15
 Arg Pro Pro Arg Gln Thr Pro Gly Thr Gln Pro Ala Ala Pro His Ala
 20 25 30
 Trp Gly Met Leu Asn Asp Met Gln Trp Leu Ala Ser Ser Asp Ser Glu
 35 40 45
 Glu Glu Thr Glu Val Gly Ile Ser Asp Asp Asp Leu His Arg Asp Ser
 50 55 60
 Thr Ser Glu Ala Gly Ser Thr Asp Thr Glu Met Phe Glu Ala Gly Leu
 65 70 75 80
 Met Asp Ala Ala Thr Pro Pro Ala Arg Pro Pro Ala Glu Arg Gln Gly
 85 90 95
 Ser Pro Thr Pro Ala Asp Ala Gln Gly Ser Cys Gly Gly Gly Pro Val
 100 105 110
 Gly Glu Glu Glu Ala Glu Ala Gly Gly Gly Gly Asp Val Cys Ala Val
 115 120 125
 Cys Thr Asp Glu Ile Ala Pro Pro Leu Arg Cys Gln Ser Phe Pro Cys
 130 135 140
 Leu His Pro Phe Cys Ile Pro Cys Met Lys Thr Trp Ile Pro Leu Arg
 145 150 155 160
 Asn Thr Cys Pro Leu Cys Asn Thr Pro Val Ala Tyr Leu Ile Val Gly
 165 170 175
 Val Thr Ala Ser Gly Ser Phe Ser Thr Ile Pro Ile Val Asn Asp Pro
 180 185 190
 Arg Thr Arg Val Glu Ala Glu Ala Ala Val Arg Ala Gly Thr Ala Val
 195 200 205
 Asp Phe Ile Trp Thr Gly Asn Pro Arg Thr Ala Pro Arg Ser Leu Ser
 210 215 220

Leu Gly Gly His Thr Val Arg Ala Leu Ser Pro Thr Pro Pro Trp Pro
 225 230 235 240
 Gly Thr Asp Asp Glu Asp Asp Asp Leu Ala Asp Gly Val Asp Tyr Val
 245 250 255
 Pro Pro Ala Pro Arg Arg Ala Pro Arg Arg Gly Gly Gly Gly Ala Gly
 260 265 270
 Ala Thr Arg Gly Thr Ser Gln Pro Ala Ala Thr Arg Pro Ala Pro Pro
 275 280 285
 Gly Ala Pro Arg Ser Ser Ser Ser Gly Gly Ala Pro Leu Arg Ala Gly
 290 295 300
 Val Gly Ser Gly Ser Gly Gly Gly Pro Ala Val Ala Val Val Pro
 305 310 315 320
 Arg Val Ala Ser Leu Pro Pro Ala Ala Gly Gly Gly Arg Ala Gln Ala
 325 330 335
 Arg Arg Val Gly Glu Asp Ala Ala Ala Ala Glu Gly Arg Thr Pro Pro
 340 345 350

 Ala Arg Gln Pro Arg Ala Ala Gln Glu Pro Pro Ile Val Ile Ser Asp
 355 360 365
 Ser Pro Pro Pro Ser Pro Arg Arg Pro Ala Gly Pro Gly Pro Leu Ser
 370 375 380
 Phe Val Ser Ser Ser Ser Ala Gln Val Ser Ser Gly Pro Gly Gly Gly
 385 390 395 400
 Gly Leu Pro Gln Ser Ser Gly Arg Ala Ala Arg Pro Arg Ala Ala Val
 405 410 415
 Ala Pro Arg Val Arg Ser Pro Pro Arg Ala Ala Ala Ala Pro Val Val
 420 425 430
 Ser Ala Ser Ala Asp Ala Ala Gly Pro Ala Pro Pro Ala Val Pro Val
 435 440 445
 Asp Ala His Arg Ala Pro Arg Ser Arg Met Thr Gln Ala Gln Thr Asp
 450 455 460
 Thr Gln Ala Gln Ser Leu Gly Arg Ala Gly Ala Thr Asp Ala Arg Gly
 465 470 475 480
 Ser Gly Gly Pro Gly Ala Glu Gly Gly Pro Gly Val Pro Arg Gly Thr
 485 490 495
 Asn Thr Pro Gly Ala Ala Pro His Ala Ala Glu Gly Ala Ala Ala Arg
 500 505 510
 Pro Arg Lys Arg Arg Gly Ser Asp Ser Gly Pro Ala Ala Ser Ser Ser
 515 520 525
 Ala Ser Ser Ser Ala Ala Pro Arg Ser Pro Leu Ala Pro Gln Gly Val
 530 535 540
 Gly Ala Lys Arg Ala Ala Pro Arg Arg Ala Pro Asp Ser Asp Ser Gly
 545 550 555 560
 Asp Arg Gly His Gly Pro Leu Ala Pro Ala Ser Ala Gly Ala Ala Pro
 565 570 575
 Pro Ser Ala Ser Pro Ser Ser Gln Ala Ala Val Ala Ala Ala Ser Ser
 580 585 590
 Ser Ser Ala Ser Ser Ser Ser Ala Ser Ser Ser Ser Ala Ser Ser Ser
 595 600 605
 Ser Ala Ser Ser Ser Ser Ala Ser Ser Ser Ser Ala Ser Ser Ser Ser
 610 615 620
 Ala Ser Ser Ser Ala Gly Gly Ala Gly Gly Ser Val Ala Ser Ala Ser
 625 630 635 640
 Gly Ala Gly Glu Arg Glu Thr Ser Leu Gly Pro Arg Ala Ala Ala
 645 650 655
 Pro Arg Gly Pro Arg Lys Cys Ala Arg Lys Thr Arg His Ala Glu Gly
 660 665 670
 Gly Pro Glu Pro Gly Ala Arg Asp Pro Ala Pro Gly Leu Thr Arg Tyr

675	680	685
Leu Pro Ile Ala Gly Val Ser Ser Val Val Ala Leu Ala Pro Tyr Val		
690	695	700
Asn Lys Thr Val Thr Gly Asp Cys Leu Pro Val Leu Asp Met Glu Thr		
705	710	715
Gly His Ile Gly Ala Tyr Val Val Leu Val Asp Gln Thr Gly Asn Val		
725	730	735
Ala Asp Leu Leu Arg Ala Ala Ala Pro Ala Trp Ser Arg Arg Thr Leu		
740	745	750
Leu Pro Glu His Ala Arg Asn Cys Val Arg Pro Pro Asp Tyr Pro Thr		
755	760	765
Pro Pro Ala Ser Glu Trp Asn Ser Leu Trp Met Thr Pro Val Gly Asn		
770	775	780
Met Leu Phe Asp Gln Gly Thr Leu Val Gly Ala Leu Asp Phe His Gly		
785	790	795
Leu Arg Ser Arg His Pro Trp Ser Arg Glu Gln Gly Ala Pro Ala Pro		
805	810	815
Ala Gly Asp Ala Pro Ala Gly His Gly Glu		
820	825	

<210> 48
 <211> 3350
 <212> DNA
 <213> HSV-2

<220>

<221> misc_feature

<222> 1027, 1034, 1054, 1055, 1056, 1057, 1058, 1059, 1060, 1061,
 1062, 1063, 1064, 1065, 1066, 1067, 1068, 1069, 1070, 1071,
 1072, 1073, 1074, 1075, 1076, 1077, 1078, 1079, 1080, 1081,
 1082, 1083, 1084, 1085, 1086, 1087, 1088, 1089, 1090

<223> n = A,T,C or G

<220>

<221> misc_feature

<222> 1091, 1092, 1093, 1094, 1095, 1096, 1097, 1098, 1099, 1100,
 1101, 1102, 1103, 1104, 1105, 1106, 1107, 1108, 1109, 1110,
 1111, 1112, 1113, 1114, 1115, 1116, 1117, 1118, 1119, 1120,
 1121, 1122, 1123, 1124, 1125, 1126, 1127, 1128, 1129

<223> n = A,T,C or G

<220>

<221> misc_feature

<222> 1130, 1131, 1132, 1133, 1134, 1135, 1136, 1137, 1138, 1139,
 1140, 1141, 1142, 1143, 1144, 1145, 1146, 1147, 1148, 1149,
 1150, 1151, 1152, 1327, 1364, 1390, 1392

<223> n = A,T,C or G

<400> 48

```

ccgtcggtga cctgcaggag ctcgtttatt aatagccagt ccatgctcag cgtagcggcc 60
agcccctggg gagacaggtc caaggagtc ggaaccaccg tcggctgacc caggggcccc 120
aggetgtagt cccccaggc cccaggtca tgacggttcg tgagcacgac gaggtctgcg 180
gccgggctgg ggggcgcgtc ctcggtcgcg tgggccatca cctcctgaat ggctgcgggtg 240
cgctgatcgg ccgagctggc gaagggcgcc acgaccagcg cgcgctccgt ctgcaggccc 300
ttccacgtgt cgtggagttc ctgaacgaac tcggccaccc gctcggggcc cgtggccgcg 360
cgcgcggcct gatagccggc cgagaggcgc cgccagcgcg ccaggaactg actcatgtaa 420
cagaaccgga ggacctggc ccccgacatc aactttgacg ccctggcggtg gatgcccgcg 480
acgatggcca ggaaccggtg gatttccgcg cgcacgacgg ccagcacgtt accctcgtgc 540

```



```

gagacctggg ccgccagctc gtcgcatacc ccgaggtgcg ccgtcgtctc ggtgacgacg 600
gaccgcagcc ccgcgagggg cgcgaccagc gcgcgcttgg cgtcgtgata catgccgcag 660
tactggctca ccgcgtcgcc catggcctcg gggcgccagg gccccaggcg ctcgtgggcg 720
tctgcgacca cggcgtagag gcggtgcccg tcgctctcga accggcactc aaagaaggcg 780
gcgagcgtgc gcatgtgaag ccgcagcagc acgatcgcggt cctccagctg gcggaccagg 840
gggtcggcgc gctcggcgag ctccctgcagc acccccgggg ccgccagggg gtacatgctg 900
atcagcagca ggtgtgtgcc cacctcgggg ggctgggggg gaggcagctg gaccgcgggc 960
cgcagctgct cgacggcccc cctggcgatc acgtacagct cgcgcagcag ctgctcgatg 1020
ttgtcnggcc atcntgcate gtgggccga cgcnnnnnnn nnnnnnnnnn nnnnnnnnnn 1080
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 1140
nnnnnnnnnn nnaggccagc accgcagggg caaactcgat ggggcggggc aggtaggcag 1200
cgttgacagt ggccctcagc gcgtcccga ccaccagggg cagcacgtaa gggacgaacc 1260
ccgggtcggc gaggacgttg ggggtgatgc cctccagggg cgggaagcgg atcttggtgg 1320
ccgcggnacg gtgaaccgag ggggcgtggc taggcggccc gacngggagc atcgcggaca 1380
gcggcgtggn cnggggtggtg ggggtcaggt ccagtggggt ctggccgtac acgtcgagcc 1440
agatgagcgc cgtctcgcg cgtctcgcg aggaggtgg gctggccggc gctgaagcgg cgtcggcccg 1500
tctcaaactc cccacagagc gtgcgccga ggctcgccag gtgttcctgc ggcacggccg 1560
ggcccatgat gcgcgccagc gtctggctga ggacgcggcc cgacaggccg accgcctcac 1620
agagccgccc gtgctgtgtc tcgctggcgc cctggatccg ccggaacgtt ttcacgtagc 1680
cggcgtagtg cccgtactcc cgcgcgagcc cgaacacgtt cgcccccgca agggcaatgc 1740
acccaaagag ctgctggatc tcgctgagcc cgtggccggg gggcgtccgc gcgggcaccc 1800
ccgccaccaa aaacccctcc agggccgata tgtactgggt gcagtgcgcg ggcgtgaacc 1860
ccgcgtcggg aagcgtgttg atcaccacgg aggcgaggtt gctgttctgg accaaagccc 1920
acgtctgctg cagcagcgcg aggagccgtt gctgggcccc ggcgaggggc ggctccccta 1980
gctgcagcag gccggtgacg gccggacgga agatggccag cgcgcagcga ctcagaaacg 2040
gcacgtcggg gtcgaagac gccgcgtccg tccgcacgcg cgccatcagc gtccccgggg 2100
gcgcgcacgc cgaccgcggg ctgacgcggc ttaggcggt cgacacgcgc acctcctcgc 2160
gactgcgaac cattttggtg gcctcgaggg gcgggatcat gatagccggg tcgatctccc 2220
gcaccgtgtg ctgaaactgg gccagcagcg gcggcgggac caccgcggcc cgatcggggg 2280
tcgtcaggta ctgctccacc agcgccagcg taaacagggg ccgcgtgagg ggggtcaggg 2340
cggcgtcgtc gatgcgtgt aggtgcgcgg agaacagcgt cacccaattg ctgaccaggg 2400
ccaagaaccg gagaccctct tgcacgatcg gggacgggaa gagcaggctg tacgccgggg 2460
tggtcagggt ggcccggggt tgcgccaggg gaaccgggga catcttaagc gacatctccc 2520
cgagggcctc cagggaggtc cgcgggttca tggccaggga gctctgggtg acggctccgc 2580
agcggctgat ccactccacg gcacactggc ggacgcgcac cggccccagg gccgcgtgg 2640
tgccgagccc ggcgccctcc agecgtggg tcgtgtcgga gccggtgatc gccaggaccg 2700
tgtccttgat gacgtccatc tcccggaagg ccgcctcggg ggtctcggg agcgccaccg 2760
ccatgcggtg caccagcagc ccggggaggt tctcggccaa gagcgccgtc tccggaagcc 2820
cgtgggcccc gtgcaaggcg cacagttgct ccaggagcgg gtgccagcac gcccgccct 2880
ccgcggggcc gaccgcgcgc cccgacaaca gaaacgcggc cgtggcgggc cgcagtttg 2940
ccgcggacag aaacgcgggc tcgtccgcgc tgcccgcgg ctcgctcgag ggggaggggc 3000
gccggcgagg gttggtcagg ctcccaaca ggacctgaa cggctcggtt gggggtggag 3060
cggacggggg ggtcatgccg gcgggcgcgg ggacctggag cgcgtgtoc gacatggcga 3120
ccggcgtgcg cgctcgcgca cgcggcgcg agaccgcgg cccaaacggg aatgactgcc 3180
gccgccctat acggaggggc taagtatcg ccggggaccc ttcgaaacc cgggcgtgtc 3240
gcaagtacgc cgcgaaggcg cggcgtgtta tacggcgct tatgtcccgg cattccgttc 3300
gtgggttcgg gcccggtgc tgtcgggtgg gagtgtgtgt gtgtggggg 3350

```

<210> 49

<211> 3345

<212> DNA

<213> HSV-2

<400> 49

```

atgtcggaca gcgcgtccca ggtcccggcg cccgcgggca tgaccccccc gtccgtccca 60
ccccaaacg gaccgttgca ggtcctgttg gggagcctga ccaacctccg ccggccgccc 120
tccccctcga gcgagccggc gggcagcgcg gacgagccgg cgtttctgtc cgcggccaaa 180
ctgcacgccg ccacggcggc gtttctgttg tcgggcgcgg cggtcggccc ggcgaggcgc 240

```

```

cgggcggtgct ggcacccgct cctggagcaa ctgtgcgct tgcaccgggc ccaagggtt 300
ccggagacgg cgctcttggc cgagaacctc cccgggctgc tgggtgaccg catggcggtg 360
gcgtcccccg agacccccga ggcggccttc cgggagatgg acgtcatcaa ggacacgggtc 420
ctggcgatca ccggctccga cacgacccac gcgtggagg ccgcccgggt gcgcaccacg 480
gcggccctgg ggccgggtgcg cgtccgccag tgtgccgtgg agtggatcga ccgtggcg 540
accgtcaccc agagctgcct ggccatgaac ccgcgaccc cctggaggc cctcggggag 600
atgtcgctta agatgtcccc ggttccctg gggcaacccg gcgccaacct gaccaccccg 660
gcgtacagcc tgctcttccc gtcccgatc gtgcaagagg gtctccggtt cttggccctg 720
gtcagcaatt ggggtgacgt gttctcggcg cactacagc gcatcgacga cgcgcctt 780
acccccctca cgcgggcccct gtttacgtg agcgtgggtg acgactacct gacgacccc 840
gatcgggggc cggtgggtccc gccgcgctg ctggcccagt ttcagcacac ggtgcgggag 900
atcgacccgg ctatcatgat cccgcccctc gaggccacca aaatggttcg cagtgcgag 960
gaggtgcgcg tgtcgaccgc cctaagccgc gtcagcccgc ggtcgcggtg cgcgcccccg 1020
gggacgctga tggcgcgct gcggacggac ggcggcgtct tcgacccga cgtgccgttt 1080
ctgagtgcgt cggcgctggc catcttccgt ccggcgtca ccggcctgct gcagctagg 1140
gagccgccc cgcgggggc ccagcaacgg etcctcgcg tgctgcagca gacgtgggt 1200
ttggtccaga acagcaatc gccctcgtg gtgatcaaca cgcttaccga cgcggggtt 1260
acgcccgcgc actgcacca gtacatatcg gccctggagg ggttttggg ggcgggggtg 1320
ccgcgcgga cgcccccg ccacgggctc agcgagatcc agcagctct tgggtgcatt 1380
gcccttgcg gggcgaaagt gttcgggct gcgcgggagt acgggacta cgcgggtac 1440
gtgaaaacgt tccggcggt ccaggcgcc cagcagcaca cgcacgggcg gctctgtgag 1500
gcggtcggcc tgtcgggcgg cgtcctcagc cagacgctgg cgcgcatcat gggcccggcc 1560
gtgcgacgg aacacctggc gagcctgcgg cgcacgctcg tgggggagtt tgagacggcc 1620
gagcgccgct tcagcgccgg ccagcccagc etcctgcgag agacggcgt catctggctc 1680
gacgtgtacg gccagacca ctgggacctg acccccacca ccccgccac gccgtgtcc 1740
gcgtgetcc cgtcgggcc gcctagccac gcccctcgg ttacctggc cgcggccacc 1800
aagatccgct tcccggcct ggagggcac caccccaacg tcctcgcca cccggggtt 1860
gtcccttacg tgctggcct ggtggtcggg gacgcgtga gggccacgtg caacgctgcc 1920
tacctgcccc gcccatcga gtttgcctg cgggtgctgg cctgggcgag cgacttcggc 1980
ctgggctatc tcccaccgt cgaggggac cgcacaaat tgggcgcgt gatcacctc 2040
ctgaacgg ccaccgggc cggcgtcggg ccacgatgc agatggcga caacatcgag 2100
cagctgtcgc cgtgatcgc agggggcg tcgagcagct ggcggccgag 2160
gtccagctgc ctcccccca gcctccgag gtgggcagca gcctgctgt gatcagcatg 2220
tacgccttg cggccgggg ggtgctgag gagctcgcc agcgcgcga cccctggtc 2280
cgccagctg aggacgcgat cgtgctgtg cggctgcaca tgcgcagct cgcgccttc 2340
tttgagtgc ggttcgagag cgacgggac cgcctgtac ccgtggtcgc agacgccac 2400
gagcgccgt ggccctggc ccccgaggcc atgggcgag cggtgagcca gtactgcggc 2460
atgtatcac acgccaagc cgcgtggtc gcgtccctc cggggtgag cgtccgtcgtc 2520
accgagacga cggcgacct cggggtatgc gacgagctg cggcccagg ctgcgacgag 2580
ggtaacgtgc tggccgtcgt gcggcgggaa atccacgggt tcctggccat cgtgtcgggc 2640
atccacgcca gggcgtcaaa gttgatgtc ggggaccagg tccccgggt ctgttacatg 2700
agtcaattc tggcgcgct gcggcgctc tcggccggct acaggccgc acgcgcggcc 2760
acgggcccc agcggggtgc cagttcgtt caggaaactc acgacagtg gaaggcctg 2820
cagacggagc gcgcgtggt cgtggcggc ttcgccagct cggccgatca gcgcaccga 2880
gccattcagg aggtgatgg ccacgcgacc gaggacgcgc ccccgagcc ggccgagac 2940
ctcgtcgtc tcacgaacc tcatgacct ggggcctgg gggactacag cctggggccc 3000
ctgggtcagc cgacggtggt tccggactc gtggacctg ctcccaggg gctggccgct 3060
acgtgagca tggactggt attaataaac gagctcctg aggtcaccga cggcgtgtt 3120
cgcgcctcg cgtttcgcc ttcgcggc ccggggggcc cgggggacct ggaggccaa 3180
gatgccggc gtagcacc cgaaccacg acacccggc cacaggacac gcaggccgc 3240
gcgcgctga cgcgccggc gggccgcgag acggtccct ggccaacac cccgtggag 3300
gacgacgaga tgacgccga ggagacacca ccggtacac cgtag 3345

```

```

<210> 50
<211> 993
<212> PRT
<213> HSV-2

```

<400> 50
 Glu Pro Ala Gly Ser Ala Asp Glu Pro Ala Phe Leu Ser Ala Ala Lys
 5 10 15
 Leu His Ala Ala Thr Ala Ala Phe Leu Leu Ser Gly Ala Ala Val Gly
 20 25 30
 Pro Ala Glu Ala Arg Ala Cys Trp His Pro Leu Leu Glu Gln Leu Cys
 35 40 45
 Ala Leu His Arg Ala His Gly Leu Pro Glu Thr Ala Leu Leu Ala Glu
 50 55 60
 Asn Leu Pro Gly Leu Leu Val His Arg Met Ala Val Ala Leu Pro Glu
 65 70 75 80
 Thr Pro Glu Ala Ala Phe Arg Glu Met Asp Val Ile Lys Asp Thr Val
 85 90 95
 Leu Ala Ile Thr Gly Ser Asp Thr Thr His Ala Leu Glu Ala Ala Gly
 100 105 110
 Leu Arg Thr Thr Ala Ala Leu Gly Pro Val Arg Val Arg Gln Cys Ala
 115 120 125
 Val Glu Trp Ile Asp Arg Trp Arg Thr Val Thr Gln Ser Cys Leu Ala
 130 135 140
 Met Asn Pro Arg Thr Ser Leu Glu Ala Leu Gly Glu Met Ser Leu Lys
 145 150 155 160
 Met Ser Pro Val Pro Leu Gly Gln Pro Gly Ala Asn Leu Thr Thr Pro
 165 170 175
 Ala Tyr Ser Leu Leu Phe Pro Ser Pro Ile Val Gln Glu Gly Leu Arg
 180 185 190
 Phe Leu Ala Leu Val Ser Asn Trp Val Thr Leu Phe Ser Ala His Leu
 195 200 205
 Gln Arg Ile Asp Asp Ala Ala Leu Thr Pro Leu Thr Arg Ala Leu Phe
 210 215 220
 Thr Leu Ala Leu Val Asp Asp Tyr Leu Thr Thr Pro Asp Arg Gly Ala
 225 230 235 240
 Val Val Pro Pro Pro Leu Leu Ala Gln Phe Gln His Thr Val Arg Glu
 245 250 255
 Ile Asp Pro Ala Ile Met Ile Pro Pro Leu Glu Ala Thr Lys Met Val
 260 265 270
 Arg Ser Arg Glu Glu Val Arg Val Ser Thr Ala Leu Ser Arg Val Ser
 275 280 285
 Pro Arg Ser Ala Cys Ala Pro Pro Gly Thr Leu Met Ala Arg Val Arg
 290 295 300
 Thr Asp Ala Ala Val Phe Asp Pro Asp Val Pro Phe Leu Ser Ala Ser
 305 310 315 320
 Ala Leu Ala Ile Phe Arg Pro Ala Val Thr Gly Leu Leu Gln Leu Gly
 325 330 335
 Glu Pro Pro Ser Ala Gly Ala Gln Gln Arg Leu Leu Ala Leu Leu Gln
 340 345 350
 Gln Thr Trp Ala Leu Val Gln Asn Ser Asn Ser Pro Ser Val Val Ile
 355 360 365
 Asn Thr Leu Thr Asp Ala Gly Phe Thr Pro Ala His Cys Thr Gln Tyr
 370 375 380
 Ile Ser Ala Leu Glu Gly Phe Leu Val Ala Gly Val Pro Ala Arg Thr
 385 390 395 400
 Pro Pro Gly His Gly Leu Ser Glu Ile Gln Gln Leu Phe Gly Cys Ile
 405 410 415
 Ala Leu Ala Gly Ala Asn Val Phe Gly Leu Ala Arg Glu Tyr Gly His
 420 425 430
 Tyr Ala Gly Tyr Val Lys Thr Phe Arg Arg Ile Gln Gly Ala Ser Glu
 435 440 445
 His Thr His Gly Arg Leu Cys Glu Ala Val Gly Leu Ser Gly Gly Val

450		455		460
Leu Ser Gln Thr Leu	Ala Arg Ile Met Gly Pro Ala Val Pro Thr Glu			
465		470		480
His Leu Ala Ser Leu	Arg Arg Thr Leu Val Gly Glu Phe Glu Thr Ala			
	485		490	495
Glu Arg Arg Phe Ser	Ala Gly Gln Pro Ser Leu Leu Arg Glu Thr Ala			
	500		505	510
Leu Ile Trp Leu Asp	Val Tyr Gly Gln Thr His Trp Asp Leu Thr Pro			
	515		520	525
Thr Thr Pro Ala Thr	Pro Leu Ser Ala Leu Leu Pro Val Gly Pro Pro			
	530		535	540
Ser His Ala Pro Ser	Val His Leu Ala Ala Ala Thr Lys Ile Arg Phe			
545		550		560
Pro Ala Leu Glu Gly	Ile His Pro Asn Val Leu Ala Asp Pro Gly Phe			
	565		570	575
Val Pro Tyr Val Leu	Ala Leu Val Val Gly Asp Ala Leu Arg Ala Thr			
	580		585	590
Cys Asn Ala Ala Tyr	Leu Pro Arg Pro Ile Glu Phe Ala Leu Arg Val			
	595		600	605
Leu Ala Trp Ala Arg	Asp Phe Gly Leu Gly Tyr Leu Pro Thr Val Glu			
	610		615	620
Gly His Arg Thr Lys	Leu Gly Ala Leu Ile Thr Leu Leu Glu Pro Ala			
625		630		640
Thr Arg Ala Gly Val	Gly Pro Thr Met Gln Met Ala Asp Asn Ile Glu			
	645		650	655
Gln Leu Leu Arg Glu	Leu Tyr Val Ile Ala Arg Gly Ala Val Glu Gln			
	660		665	670
Leu Arg Pro Ala Val	Gln Leu Pro Pro Pro Gln Pro Pro Glu Val Gly			
	675		680	685
Ser Ser Leu Leu Leu	Ile Ser Met Tyr Ala Leu Ala Ala Arg Gly Val			
	690		695	700
Leu Gln Glu Leu Ala	Glu Arg Ala Asp Pro Leu Val Arg Gln Leu Glu			
705		710		720
Asp Ala Ile Val Leu	Leu Arg Leu His Met Arg Thr Leu Ala Ala Phe			
	725		730	735
Phe Glu Cys Arg Phe	Glu Ser Asp Gly His Arg Leu Tyr Ala Val Val			
	740		745	750
Ala Asp Ala His Glu	Arg Leu Gly Pro Trp Arg Pro Glu Ala Met Gly			
	755		760	765
Asp Ala Val Ser Gln	Tyr Cys Gly Met Tyr His Asp Ala Lys Arg Ala			
	770		775	780
Leu Val Ala Ser Leu	Ala Gly Leu Arg Ser Val Val Thr Glu Thr Thr			
785		790		800
Ala His Leu Gly Val	Cys Asp Glu Leu Ala Ala Gln Val Ser His Glu			
	805		810	815
Gly Asn Val Leu Ala	Val Val Arg Arg Glu Ile His Gly Phe Leu Ala			
	820		825	830
Ile Val Ser Gly Ile	His Ala Arg Ala Ser Lys Leu Met Ser Gly Asp			
	835		840	845
Gln Val Pro Gly Phe	Cys Tyr Met Ser Gln Phe Leu Ala Arg Trp Arg			
	850		855	860
Arg Leu Ser Ala Gly	Tyr Gln Ala Ala Arg Ala Ala Thr Gly Pro Glu			
865		870		880
Arg Val Ala Glu Phe	Val Gln Glu Leu His Asp Thr Trp Lys Gly Leu			
	885		890	895
Gln Thr Glu Arg Ala	Leu Val Val Ala Arg Phe Ala Ser Ser Ala Asp			
	900		905	910
Gln Arg Thr Ala Ala	Ile Gln Glu Val Met Ala His Ala Thr Glu Asp			

<400> 51	Met	Ser	Asp	Ser	Ala	Leu	Gln	Val	Pro	Ala	Pro	Ala	Gly	Met	Thr	Pro
				5						10					15	
Pro	Ser	Ala	Pro	Pro	Pro	Asn	Gly	Pro	Leu	Gln	Val	Leu	Leu	Gly	Ser	
			20					25						30		
Leu	Thr	Asn	Leu	Arg	Arg	Pro	Pro	Ser	Pro	Ser	Ser	Glu	Pro	Ala	Gly	
		35					40					45				
Ser	Ala	Asp	Glu	Pro	Ala	Phe	Leu	Ser	Ala	Ala	Lys	Leu	His	Ala	Ala	
	50					55					60					
Thr	Ala	Ala	Phe	Leu	Leu	Ser	Gly	Ala	Ala	Val	Gly	Pro	Ala	Glu	Ala	
65					70					75					80	
Arg	Ala	Cys	Trp	His	Pro	Leu	Leu	Glu	Gln	Leu	Cys	Ala	Leu	His	Arg	
				85					90					95		
Ala	His	Gly	Leu	Pro	Glu	Thr	Ala	Leu	Leu	Ala	Glu	Asn	Leu	Pro	Gly	
			100					105					110			
Leu	Leu	Val	His	Arg	Met	Ala	Val	Ala	Leu	Pro	Glu	Thr	Pro	Glu	Ala	
		115					120					125				
Ala	Phe	Arg	Glu	Met	Asp	Val	Ile	Lys	Asp	Thr	Val	Leu	Ala	Ile	Thr	
	130					135					140					
Gly	Ser	Asp	Thr	Thr	His	Ala	Leu	Glu	Ala	Ala	Gly	Leu	Arg	Thr	Thr	
145					150					155					160	
Ala	Ala	Leu	Gly	Pro	Val	Arg	Val	Arg	Gln	Cys	Ala	Val	Glu	Trp	Ile	
				165					170					175		
Asp	Arg	Trp	Arg	Thr	Val	Thr	Gln	Ser	Cys	Leu	Ala	Met	Asn	Pro	Arg	
			180					185					190			
Thr	Ser	Leu	Glu	Ala	Leu	Gly	Glu	Met	Ser	Leu	Lys	Met	Ser	Pro	Val	
		195					200					205				
Pro	Leu	Gly	Gln	Pro	Gly	Ala	Asn	Leu	Thr	Thr	Pro	Ala	Tyr	Ser	Leu	
	210					215					220					
Leu	Phe	Pro	Ser	Pro	Ile	Val	Gln	Glu	Gly	Leu	Arg	Phe	Leu	Ala	Leu	
225					230					235					240	
Val	Ser	Asn	Trp	Val	Thr	Leu	Phe	Ser	Ala	His	Leu	Gln	Arg	Ile	Asp	
				245					250					255		
Asp	Ala	Ala	Leu	Thr	Pro	Leu	Thr	Arg	Ala	Leu	Phe	Thr	Leu	Ala	Leu	
			260					265					270			
Val	Asp	Asp	Tyr	Leu	Thr	Thr	Pro	Asp	Arg	Gly	Ala	Val	Val	Pro	Pro	
		275					280					285				
Pro	Leu	Leu	Ala	Gln	Phe	Gln	His	Thr	Val	Arg	Glu	Ile	Asp	Pro	Ala	
	290					295					300					
Ile	Met	Ile	Pro	Pro	Leu	Glu	Ala	Thr	Lys	Met	Val	Arg	Ser	Arg	Glu	
305					310					315					320	
Glu	Val	Arg	Val	Ser	Thr	Ala										

325 330 335
 Cys Ala Pro Pro Gly Thr Leu Met Ala Arg Val Arg Thr Asp Ala Ala
 340 345 350
 Val Phe Asp Pro Asp Val Pro Phe Leu Ser Ala Ser Ala Leu Ala Ile
 355 360 365
 Phe Arg Pro Ala Val Thr Gly Leu Leu Gln Leu Gly Glu Pro Pro Ser
 370 375 380
 Ala Gly Ala Gln Gln Arg Leu Leu Ala Leu Leu Gln Gln Thr Trp Ala
 385 390 395 400
 Leu Val Gln Asn Ser Asn Ser Pro Ser Val Val Ile Asn Thr Leu Thr
 405 410 415
 Asp Ala Gly Phe Thr Pro Ala His Cys Thr Gln Tyr Ile Ser Ala Leu
 420 425 430
 Glu Gly Phe Leu Val Ala Gly Val Pro Ala Arg Thr Pro Pro Gly His
 435 440 445
 Gly Leu Ser Glu Ile Gln Gln Leu Phe Gly Cys Ile Ala Leu Ala Gly
 450 455 460
 Ala Asn Val Phe Gly Leu Ala Arg Glu Tyr Gly His Tyr Ala Gly Tyr
 465 470 475 480
 Val Lys Thr Phe Arg Arg Ile Gln Gly Ala Ser Glu His Thr His Gly
 485 490 495
 Arg Leu Cys Glu Ala Val Gly Leu Ser Gly Gly Val Leu Ser Gln Thr
 500 505 510
 Leu Ala Arg Ile Met Gly Pro Ala Val Pro Thr Glu His Leu Ala Ser
 515 520 525
 Leu Arg Arg Thr Leu Val Gly Glu Phe Glu Thr Ala Glu Arg Arg Phe
 530 535 540
 Ser Ala Gly Gln Pro Ser Leu Leu Arg Glu Thr Ala Leu Ile Trp Leu
 545 550 555 560
 Asp Val Tyr Gly Gln Thr His Trp Asp Leu Thr Pro Thr Thr Pro Ala
 565 570 575
 Thr Pro Leu Ser Ala Leu Leu Pro Val Gly Pro Pro Ser His Ala Pro
 580 585 590
 Ser Val His Leu Ala Ala Ala Thr Lys Ile Arg Phe Pro Ala Leu Glu
 595 600 605
 Gly Ile His Pro Asn Val Leu Ala Asp Pro Gly Phe Val Pro Tyr Val
 610 615 620
 Leu Ala Leu Val Val Gly Asp Ala Leu Arg Ala Thr Cys Asn Ala Ala
 625 630 635 640
 Tyr Leu Pro Arg Pro Ile Glu Phe Ala Leu Arg Val Leu Ala Trp Ala
 645 650 655
 Arg Asp Phe Gly Leu Gly Tyr Leu Pro Thr Val Glu Gly His Arg Thr
 660 665 670
 Lys Leu Gly Ala Leu Ile Thr Leu Leu Glu Pro Ala Thr Arg Ala Gly
 675 680 685
 Val Gly Pro Thr Met Gln Met Ala Asp Asn Ile Glu Gln Leu Leu Arg
 690 695 700
 Glu Leu Tyr Val Ile Ala Arg Gly Ala Val Glu Gln Leu Arg Pro Ala
 705 710 715 720
 Val Gln Leu Pro Pro Pro Gln Pro Pro Glu Val Gly Ser Ser Leu Leu
 725 730 735
 Leu Ile Ser Met Tyr Ala Leu Ala Ala Arg Gly Val Leu Gln Glu Leu
 740 745 750
 Ala Glu Arg Ala Asp Pro Leu Val Arg Gln Leu Glu Asp Ala Ile Val
 755 760 765
 Leu Leu Arg Leu His Met Arg Thr Leu Ala Ala Phe Phe Glu Cys Arg
 770 775 780
 Phe Glu Ser Asp Gly His Arg Leu Tyr Ala Val Val Ala Asp Ala His

785		790		795		800
Glu Arg Leu Gly Pro	Trp Arg Pro Glu Ala Met Gly Asp Ala Val Ser					
	805		810		815	
Gln Tyr Cys Gly Met Tyr His Asp Ala Lys Arg Ala Leu Val Ala Ser						
	820		825		830	
Leu Ala Gly Leu Arg Ser Val Val Thr Glu Thr Thr Ala His Leu Gly						
	835		840		845	
Val Cys Asp Glu Leu Ala Ala Gln Val Ser His Glu Gly Asn Val Leu						
	850		855		860	
Ala Val Val Arg Arg Glu Ile His Gly Phe Leu Ala Ile Val Ser Gly						
	865		870		875	880
Ile His Ala Arg Ala Ser Lys Leu Met Ser Gly Asp Gln Val Pro Gly						
	885		890		895	
Phe Cys Tyr Met Ser Gln Phe Leu Ala Arg Trp Arg Arg Leu Ser Ala						
	900		905		910	
Gly Tyr Gln Ala Ala Arg Ala Ala Thr Gly Pro Glu Arg Val Ala Glu						
	915		920		925	
Phe Val Gln Glu Leu His Asp Thr Trp Lys Gly Leu Gln Thr Glu Arg						
	930		935		940	
Ala Leu Val Val Ala Arg Phe Ala Ser Ser Ala Asp Gln Arg Thr Ala						
	945		950		955	960
Ala Ile Gln Glu Val Met Ala His Ala Thr Glu Asp Ala Pro Pro Ser						
	965		970		975	
Pro Ala Ala Asp Leu Val Val Leu Thr Asn Arg His Asp Leu Gly Ala						
	980		985		990	
Trp Gly Asp Tyr Ser Leu Gly Pro Leu Gly Gln Pro Thr Val Val Pro						
	995		1000		1005	
Asp Ser Val Asp Leu Ser Pro Gln Gly Leu Ala Ala Thr Leu Ser Met						
	1010		1015		1020	
Asp Trp Leu Leu Ile Asn Glu Leu Leu Gln Val Thr Asp Gly Val Phe						
	1025		1030		1035	1040
Arg Ala Ser Ala Phe Arg Pro Ser Ala Gly Pro Gly Ala Pro Gly Asp						
	1045		1050		1055	
Leu Glu Ala Gln Asp Ala Gly Gly Ser Thr Pro Glu Pro Thr Thr Pro						
	1060		1065		1070	
Gly Pro Gln Asp Thr Gln Ala Arg Ala Pro Ser Thr Pro Ala Gly Arg						
	1075		1080		1085	
Glu Thr Val Pro Trp Pro Asn Thr Pro Val Glu Asp Asp Glu Met Thr						
	1090		1095		1100	
Pro Gln Glu Thr Pro Pro Val His Pro						
1105		1110				

<210> 52
 <211> 3113
 <212> DNA
 <213> HSV-2

<400> 52
 atgtcggaca ggcgcgtcca ggtcccggcg cccgccggca tgaccccccc gtcgcgtcca 60
 ccccaaacg gaccgttgca ggtcctgttg gggagcctga ccaacctccg ccggccgccc 120
 tccccctcga gcgagccggc gggcagcgcg gacgagccgg cgtttctgtc cgcggccaaa 180
 ctgcgcgccg ccacggcggc gtttctgttg tcgggcgcgg cggtcggccc ggccgagggc 240
 cgggcgtgct ggcacccgct cctggagcaa ctgtgcgcct tgcaccgggc ccacgggctt 300
 ccggagacgg cgctcttggc cgagaacctc cccgggctgc tgggtgcaccg catggcggtg 360
 cgctccccg agacccccg ggcggccttc cgggagatgg acgtcatcaa ggacacggtc 420
 ctggcgatca ccggctccga caccgaccac gcgctggagg ccgccgggct gcgcaccacg 480
 gcggccctgg ggccgggtgc cgtccgccag tgtgccgtgg agtggatcga ccgctggcgg 540

```

acggtcaccc agagctgcct ggccatgaac ccgcggaacct ccctggaggg cctcggggag 600
atgtcgctta agatgtcccc ggttccccct gggcaaccgg gcgccaacct gaccaccccg 660
gogtacagcc tgctcttccc gtccccgatc gtgcaagagg gtctccggtt cttggccctg 720
gtcagcaatt gggtagacgt gttctcgggc cacctacagc gcatcgacga cgcgcacctg 780
acccccctca cgcggggcct gtttacgctg gcgctgggtg acgagtacct gacgaccccc 840
gatcggggcg cggtagtccc gccgcgctg ctggcccagt ttcagcacac ggtgcggggag 900
atcgacccgg ctatcatgat cccgcccctc gaggccacca aaatggttcg cagtgcgcgag 960
gaggtgcgcg tgtcgaccgc cctaagccgc gtcagcccgc ggtcggcggt gcgcgccccg 1020
gggacgctga tggcgcgcggt gcggacggac gggcgccgtc tcgaccccca cgtgcccgtt 1080
ctgagtgcgt cggcgctggc catcttccgt ccggccgtca ccggcctgct gcagctaggg 1140
gagccgcccc ccgcccgggc ccagcaacgg ctctcgcgcc tgctgcagca gacgtgggct 1200
ttggtccaga acagcaactc gccctccgtg gtagtcaaca cgcttaccga cgcgggggtt 1260
acgcccgcgc actgcaccca gtacatatcg gccctggagg ggtttttggt ggcgggggtg 1320
cccgcgcgga cgcggccggg ccaagggtc agcgagatcc agcagctctt tgggtgcatt 1380
gcccttgctg gggcgaaagt gttcgggtc gcgcgggagt acgggcaacta cgcgggctac 1440
gtgaaaacgt tccggcggtt ccaggcgccc agcgagcaca cgcacggggc gctctgtgag 1500
gcggtcgggc tgtcggggcg cgtcctcagc cagacgctgg cgcgcacatc gggcccggcc 1560
gtgccgacgg aacacctggc gagcctgcgg cgcacgctcg tgggggagtt tgagacggcc 1620
gagcgccgct tcagcgccgg ccagcccagc ctctcgcgcg agacggcgct catctggctc 1680
gacgtgtacg gccagaccca ctgggacctg acccccacca ccccggccac gccgctgtcc 1740
gcgctgtccc ccgtcgggcc gcctagccac gccccctcgg ttacactggc cgcggccacc 1800
aagatccgct tcccggccct ggagggcatc caccacaacg tctcgcgga cccgggggtt 1860
gtcccttacg tgctggccct ggtggtcggg gacgcgtga gggccaagt caacgctgcc 1920
tacctgcccc gccccatcga gtttgccctg cgggtgctgg cctggggcgcg cgacttcggc 1980
ctgggctatc tccccaccgt cgaggggcac cgcacaaaat tgggcgctg gatcacctc 2040
ctcgaaccgg ccaccggggc cggcgctggg ccacgatgc agatggcca caacatcgag 2100
cagctgtgct gcgagctgta cgtgatcgcc aggggggccc tcgagcagct gcggcccgcg 2160
gtccagctgc ctcccccca gcctcccag gtgggcagca gcctgctgct gatcagcatg 2220
tacgccctgg cggcccgggg ggtgctgcag gagctcgccg agcgcgcgga cccctgggtc 2280
cgccagctgg aggacgcgat cgtgctgctg cggcttcaca tgcgcacgct cgcgccttc 2340
tttgagtgcc ggttcgagag cgacgggcac cgcctgtacg ccgtggtcgc agacggccac 2400
gagcgccctg gccctggcg ccccgaggcc atgggcgacg cggtagacca gtactcggc 2460
atgtatcacg acgccaagcg cgcgctggtc gcgtccctcg cggggtgctg gtccgtcgtc 2520
accgatcaga cggcgacct cggggtatgc acgagctgg cggcccagg ctgcacagag 2580
ggtaacgtgc tggccgctgt gccgcgggaa atccacgggt tctggccat cgtgtcgggc 2640
atccacgcca gggcgctcaa gttgatgtcg ggggaccagg tccccgggt ctgttacatg 2700
agtacgttcc tggcgcgctg gcggcgctc tcggccggt atcaggccgc gcgcgcggcc 2760
acgggccccg agcgggtggc cgagttcgtt caggaaactc acgacacgtg gaagggcctg 2820
cagacggagc gcgcgctggt cgtggcgccc ttcgccagct cggccgatca gcgcaccgca 2880
gccattcagg aggtgatggc ccacgcgacc gaggcgcgc ccccagccc ggccgcagac 2940
ctcgtcgtgc tcacgaaccg tcatgacctg ggggcctggg gggactacag cctggggccc 3000
ctgggtcagc cgacggtggg tccggactcc gtggacctgt ctcccaggg gctggccgct 3060
acgctgagca tggactgggt attaataaac gagctcctgc aggtcaccga cgg 3113

```

<210> 53

<211> 761

<212> DNA

<213> HSV-2

<400> 53

```

gcgcccgtct ggcgctcagc gcgaggccgc cgggggtttac gacgcggtgc ggacctgggg 60
gccagacgcg gaggccgaac cggaccagat ggaaaaacag tatctgctgc ccgacgatga 120
cgccgccatg cccgcggggc tcgggcttg gcacaccccc gcccgcgaca ccaccgccc 180
cgcttgccg gccgaaagcc acgcccccg cgcctccctc gaggacgcag attccattta 240
cagtcggtg agcgaggtg gggggcgct ctacgaggag atcccytggg ttcgggtata 300
cgaaaacatc tgccttcgcc ggcaagacgc cggcgggggc gcccgccgg gagacgccc 360
ggactccccg tacatcgagg cggaatcc cctgtacgac tggggcggt ctgccctctt 420

```


<400>	54																
Met	Ser	Asp	Ser	Ala	Leu	Gln	Val	Pro	Ala	Pro	Ala	Gly	Met	Thr	Pro		
				5					10					15			
Pro	Ser	Ala	Pro	Pro	Pro	Asn	Gly	Pro	Leu	Gln	Val	Leu	Leu	Gly	Ser		
			20					25					30				
Leu	Thr	Asn	Leu	Arg	Arg	Pro	Pro	Ser	Pro	Ser	Ser	Glu	Pro	Ala	Gly		
		35					40					45					
Ser	Ala	Asp	Glu	Pro	Ala	Phe	Leu	Ser	Ala	Ala	Lys	Leu	Arg	Ala	Ala		
	50					55					60						
Thr	Ala	Ala	Phe	Leu	Leu	Ser	Gly	Ala	Ala	Val	Gly	Pro	Ala	Glu	Ala		
65					70					75					80		
Arg	Ala	Cys	Trp	His	Pro	Leu	Leu	Glu	Gln	Leu	Cys	Ala	Leu	His	Arg		
				85					90					95			
Ala	His	Gly	Leu	Pro	Glu	Thr	Ala	Leu	Leu	Ala	Glu	Asn	Leu	Pro	Gly		
			100					105					110				
Leu	Leu	Val	His	Arg	Met	Ala	Val	Ala	Leu	Pro	Glu	Thr	Pro	Glu	Ala		
		115					120					125					
Ala	Phe	Arg	Glu	Met	Asp	Val	Ile	Lys	Asp	Thr	Val	Leu	Ala	Ile	Thr		
	130					135					140						
Gly	Ser	Asp	Thr	Thr	His	Ala	Leu	Glu	Ala	Ala	Gly	Leu	Arg	Thr	Thr		
145					150				155						160		
Ala	Ala	Leu	Gly	Pro	Val	Arg	Val	Arg	Gln	Cys	Ala	Val	Glu	Trp	Ile		
				165					170					175			
Asp	Arg	Trp	Arg	Thr	Val	Thr	Gln	Ser	Cys	Leu	Ala	Met	Asn	Pro	Arg		
			180					185					190				
Thr	Ser	Leu	Glu	Ala	Leu	Gly	Glu	Met	Ser	Leu	Lys	Met	Ser	Pro	Val		
		195					200					205					
Pro	Leu	Gly	Gln	Pro	Gly	Ala	Asn	Leu	Thr	Thr	Pro	Ala	Tyr	Ser	Leu		
	210					215					220						
Leu	Phe	Pro	Ser	Pro	Ile	Val	Gln	Glu	Gly	Leu	Arg	Phe	Leu	Ala	Leu		
225					230					235					240		

Val Ser Asn Trp Val Thr Leu Phe Ser Ala His Leu Gln Arg Ile Asp
 245 250 255
 Asp Ala Ala Leu Thr Pro Leu Thr Arg Ala Leu Phe Thr Leu Ala Leu
 260 265 270
 Val Asp Glu Tyr Leu Thr Thr Pro Asp Arg Gly Ala Val Val Pro Pro
 275 280 285
 Pro Leu Leu Ala Gln Phe Gln His Thr Val Arg Glu Ile Asp Pro Ala
 290 295 300
 Ile Met Ile Pro Pro Leu Glu Ala Thr Lys Met Val Arg Ser Arg Glu
 305 310 315 320
 Glu Val Arg Val Ser Thr Ala Leu Ser Arg Val Ser Pro Arg Ser Ala
 325 330 335
 Cys Ala Pro Pro Gly Thr Leu Met Ala Arg Val Arg Thr Asp Ala Ala
 340 345 350
 Val Phe Asp Pro Asp Val Pro Phe Leu Ser Ala Ser Ala Leu Ala Ile
 355 360 365
 Phe Arg Pro Ala Val Thr Gly Leu Leu Gln Leu Gly Glu Pro Pro Ser
 370 375 380
 Ala Gly Ala Gln Gln Arg Leu Leu Ala Leu Leu Gln Gln Thr Trp Ala
 385 390 395 400
 Leu Val Gln Asn Ser Asn Ser Pro Ser Val Val Ile Asn Thr Leu Thr
 405 410 415
 Asp Ala Gly Phe Thr Pro Ala His Cys Thr Gln Tyr Ile Ser Ala Leu
 420 425 430
 Glu Gly Phe Leu Val Ala Gly Val Pro Ala Arg Thr Pro Pro Gly His
 435 440 445
 Gly Leu Ser Glu Ile Gln Gln Leu Phe Gly Cys Ile Ala Leu Ala Gly
 450 455 460
 Ala Asn Val Phe Gly Leu Ala Arg Glu Tyr Gly His Tyr Ala Gly Tyr
 465 470 475 480
 Val Lys Thr Phe Arg Arg Ile Gln Gly Ala Ser Glu His Thr His Gly
 485 490 495
 Arg Leu Cys Glu Ala Val Gly Leu Ser Gly Gly Val Leu Ser Gln Thr
 500 505 510
 Leu Ala Arg Ile Met Gly Pro Ala Val Pro Thr Glu His Leu Ala Ser
 515 520 525
 Leu Arg Arg Thr Leu Val Gly Glu Phe Glu Thr Ala Glu Arg Arg Phe
 530 535 540
 Ser Ala Gly Gln Pro Ser Leu Leu Arg Glu Thr Ala Leu Ile Trp Leu

545		550		555		560
Asp Val Tyr Gly Gln Thr His Trp Asp Leu Thr Pro Thr Thr Pro Ala						
		565		570		575
Thr Pro Leu Ser Ala Leu Leu Pro Val Gly Pro Pro Ser His Ala Pro						
		580		585		590
Ser Val His Leu Ala Ala Ala Thr Lys Ile Arg Phe Pro Ala Leu Glu						
		595		600		605
Gly Ile His Pro Asn Val Leu Ala Asp Pro Gly Phe Val Pro Tyr Val						
		610		615		620
Leu Ala Leu Val Val Gly Asp Ala Leu Arg Ala Thr Cys Asn Ala Ala						
		625		630		640
Tyr Leu Pro Arg Pro Ile Glu Phe Ala Leu Arg Val Leu Ala Trp Ala						
		645		650		655
Arg Asp Phe Gly Leu Gly Tyr Leu Pro Thr Val Glu Gly His Arg Thr						
		660		665		670
Lys Leu Gly Ala Leu Ile Thr Leu Leu Glu Pro Ala Thr Arg Ala Gly						
		675		680		685
Val Gly Pro Thr Met Gln Met Ala Asp Asn Ile Glu Gln Leu Leu Arg						
		690		695		700
Glu Leu Tyr Val Ile Ala Arg Gly Ala Val Glu Gln Leu Arg Pro Ala						
		705		710		720
Val Gln Leu Pro Pro Pro Gln Pro Pro Glu Val Gly Ser Ser Leu Leu						
		725		730		735
Leu Ile Ser Met Tyr Ala Leu Ala Ala Arg Gly Val Leu Gln Glu Leu						
		740		745		750
Ala Glu Arg Ala Asp Pro Leu Val Arg Gln Leu Glu Asp Ala Ile Val						
		755		760		765
Leu Leu Arg Leu His Met Arg Thr Leu Ala Ala Phe Phe Glu Cys Arg						
		770		775		780
Phe Glu Ser Asp Gly His Arg Leu Tyr Ala Val Val Ala Asp Ala His						
		785		790		800
Glu Arg Leu Gly Pro Trp Arg Pro Glu Ala Met Gly Asp Ala Val Ser						
		805		810		815
Gln Tyr Cys Gly Met Tyr His Asp Ala Lys Arg Ala Leu Val Ala Ser						
		820		825		830
Leu Ala Gly Leu Arg Ser Val Val Thr Glu Thr Thr Ala His Leu Gly						
		835		840		845
Val Cys Asp Glu Leu Ala Ala Gln Val Ser His Glu Gly Asn Val Leu						
		850		855		860

Ala Val Val Arg Arg Glu Ile His Gly Phe Leu Ala Ile Val Ser Gly
 865 870 875 880

Ile His Ala Arg Ala Ser Lys Leu Met Ser Gly Asp Gln Val Pro Gly
 885 890 895

Phe Cys Tyr Met Ser Gln Phe Leu Ala Arg Trp Arg Arg Leu Ser Ala
 900 905 910

Gly Tyr Gln Ala Ala Arg Ala Ala Thr Gly Pro Glu Arg Val Ala Glu
 915 920 925

Phe Val Gln Glu Leu His Asp Thr Trp Lys Gly Leu Gln Thr Glu Arg
 930 935 940

Ala Leu Val Val Ala Pro Phe Ala Ser Ser Ala Asp Gln Arg Thr Ala
 945 950 955 960

Ala Ile Gln Glu Val Met Ala His Ala Thr Glu Asp Ala Pro Pro Ser
 965 970 975

Pro Ala Ala Asp Leu Val Val Leu Thr Asn Arg His Asp Leu Gly Ala
 980 985 990

Trp Gly Asp Tyr Ser Leu Gly Pro Leu Gly Gln Pro Thr Val Val Pro
 995 1000 1005

Asp Ser Val Asp Leu Ser Pro Gln Gly Leu Ala Ala Thr Leu Ser Met
 1010 1015 1020

Asp Trp Leu Leu Ile Asn Glu Leu Leu Gln Val Thr Asp
 1025 1030 1035

<210> 55
 <211> 193
 <212> PRT
 <213> HSV-2

<400> 55
 Arg Pro Leu Ala Ala Gln Arg Glu Ala Ala Gly Val Tyr Asp Ala Val
 5 10 15

Arg Thr Trp Gly Pro Asp Ala Glu Ala Glu Pro Asp Gln Met Glu Asn
 20 25 30

Thr Tyr Leu Leu Pro Asp Asp Asp Ala Ala Met Pro Ala Gly Val Gly
 35 40 45

Leu Gly Ala Thr Pro Ala Ala Asp Thr Thr Ala Ala Ala Trp Pro Ala
 50 55 60

Glu Ser His Ala Pro Arg Ala Pro Ser Glu Asp Ala Asp Ser Ile Tyr
 65 70 75 80

Glu Ser Val Ser Glu Asp Gly Gly Arg Val Tyr Glu Glu Ile Pro Trp
 85 90 95

Val Arg Val Tyr Glu Asn Ile Cys Leu Arg Arg Gln Asp Ala Gly Gly
 100 105 110

Ala Ala Pro Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala Glu
 115 120 125

Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro Pro Gly
 130 135 140

Ala Thr Arg Ala Pro Asp Pro Gly Leu Ser Leu Ser Pro Met Pro Ala
 145 150 155 160

Arg Pro Arg Thr Asn Ala Leu Ala Asn Asp Gly Pro Thr Asn Val Ala
 165 170 175

Ala Leu Ser Ala Leu Leu Thr Lys Leu Lys Arg Gly Arg His Gln Ser
 180 185 190

His

<210> 56
 <211> 15
 <212> PRT
 <213> HSV-2

<400> 56
 Ser Pro Asn Thr Asp Val Arg Met Tyr Ser Gly Lys Arg Asn Gly
 5 10 15

<210> 57
 <211> 15
 <212> PRT
 <213> HSV-2

<400> 57
 Tyr Leu Ala Ala Pro Thr Gly Ile Pro Pro Ala Phe Phe Pro Ile
 5 10 15

<210> 58
 <211> 15
 <212> PRT
 <213> HSV-2

<400> 58
 Gly Val Ala Ala Ala Thr Pro Arg Pro Asp Pro Glu Asp Gly Ala
 5 10 15

<210> 59
 <211> 15
 <212> PRT
 <213> HSV-2

<400> 59

Glu Glu Ile Pro Trp Val Arg Val Tyr Glu Asn Ile Cys Pro Arg
5 10 15

<210> 60
<211> 15
<212> PRT
<213> HSV-2

<400> 60
Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala Glu Asn Pro
5 10 15

<210> 61
<211> 15
<212> PRT
<213> HSV-2

<400> 61
Pro Asp Ser Pro Tyr Ile Glu Ala Glu Asn Pro Leu Tyr Asp Trp
5 10 15

<210> 62
<211> 15
<212> PRT
<213> HSV-2

<400> 62
Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro
5 10 15

<210> 63
<211> 15
<212> PRT
<213> HSV-2

<400> 63
Ala Ile Asp Tyr Val His Cys Glu Gly Ile Ile His Arg Asp Ile
5 10 15

<210> 64
<211> 15
<212> PRT
<213> HSV-2

<400> 64
Ala Phe Pro Val Ala Leu His Ala Val Asp Ala Pro Ser Gln Phe
5 10 15

<210> 65
<211> 3429
<212> DNA
<213> HSV-2

<400> 65

```

atggccaacc gccctgccgc atccgccctc gccggagcgc ggtctccgtc cgaacgacag 60
gaaccccggg agcccgaggt cgcgcgcctt ggcggcgacc acgtgttttg caggaaagt 120
agcggcgatga tgggtgctttc cagcgatccc cccggccccc cggcctaccg cattagcgac 180
agcagctttg ttcaatgcgg ctccaactgc agtatgataa tcgacggaga cgtggcgcg 240
ggtcattttg gtgacctcga gggcgctacg tccaacggcg ccttcgtcgc gatctcaaac 300
gtcgcagccg gcggggatgg ccgaaccgcc gtcgtggcgc tcggcggaac ctccggcccc 360
tccgcgacta catccgtggg gaccagacg tcgggggagt tcctccacgg gaacccaagg 420
accccggaac cccaaggacc ccaggctgtc ccccgcccc ctctccccc ctttccatgg 480
ggccacgagt gctgcgcccg tcgcgatgcc agggcgggcg ccgagaagga cgtcggggcc 540
gcggagtcac ggtcagacgg cccgtcgtcc gactccgaaa cggaggactc ggactcctcg 600
gacgaggata cgggttcgga gacgtgtctt cgatcctctt cgatctgggc cgcagggggc 660
actgacgacg atgacagcga ctccgactcg cggtcggacg actccgtgca gcccgcgtt 720
gtcgttcgtc gcagatggag cgacggcccc gcccccgtag cctttcccaa gccccggcgc 780
ccggcgact ccccggaaaa ccccgccctg ggcgcggcga ccgggcccgg ctcgcgcgac 840
gacccgcgcg cgtcggccga ctccgattcc gcggcccacg ccgcccacc ccaggcggac 900
gtggcgccgg ttctggacag ccagccact gtgggaacgg accccggcta cccagtcctc 960
ctagaactca cgcgcgagaa cgcggaggcg gtggcgcggt ttctggggga cgcgctcgac 1020
cgcgagcccc cgtcatgct ggagtacttc tgcgtgtagc cccgcgagga gagcaagcgc 1080
gtgccccac gaaccttcgg cagcgcccc cgctcacgg aggacgactt tgggctcctg 1140
aactacgcgc tcgctgagat gcgacgcctg tgcctggacc ttccccgggt ccccccaac 1200
gcatacacgc cctatcatct gagggagtat gcgacggcgc tggttaacgg gttcaaacc 1260
ctgggtgcggc ggtccgcccg cctgtatcgc atcctggggg ttctgggcca cctgcgcac 1320
cgtaaccggg aggcctcctt tgaggaatgg atgcgtcca aggaggtgga cctggacttc 1380
gggctgacgg aaaggcttcg cgaacacgag gccagctaa tgatcctggc ccaggccctg 1440
aacccctacg actgtctgat ccacagcacc ccgaacacgc tcgtcgagcg ggggctgcag 1500
tcggcgctga agtacgaaga gttttacctc aagcgttcg gcgggacta catggagtcc 1560
gtcttccaga tgtacacccg catcgccggg tttctggcgt gccgggcgac ccgcggcatg 1620
cgccacatcg ccttggggcg acaggggtcg tgggaggaaa tgttcaagtt cttttccac 1680
cgctctacg accaccagat cgtccagctg ctacctgcta aaccccgcca tgctgaacct cggaaaccgc 1740
aactactaca cgtccagctg ctacctgcta aaccccgcca tgctgaacct cggaaaccgc 1800
ctccgggcca tcaccggcaa cgtgagcgcc atcctcgccc gcaacggggg catcgggctg 1860
tgcatgcagg cgttcaacga cgcagcccc ggcaaccgca gcatcatgcc ggccctgaag 1920
gtcctcgact ccttgggtggc ggcgcacaac aaacagagca cgcgccccac cggggcgtgc 1980
gtgtacctgg aacctggca cagcgacgtt cgggcccgtg tcagaatgaa gggcgctcct 2040
gccggcgagg aggccagcg ctgcgacaac atcttcagcg ccctctggat gccggacctg 2100
ttcttcaagc gcctgatccg ccacctcgac ggcgagaaaa acgtcacctg gtccctgttc 2160
gaccgggaca ccagcatgtc gctcgccgac tttcacggcg aggagttcga gaagctgtac 2220
gagcacctcg aggccatggg gttcggcgaa acgatcccca tcaggacct ggcgtagccc 2280
atcgtgcgca gcggggccac caccggaagc cccttcatca tgtttaagga cgcggtaaac 2340
cgccactaca tctacgacac gcaagggcg gcatccagt ggggtctgca acctgggaag cgtgaatctg 2400
atcgtccacc cggcctccaa gcgatccagt tttgggcggc tcgcgacgc cgtgcaggcg 2460
gcccgatcg tctccaggca gacgtttgac tttgggcggc tcgcgacgc cgtgcaggcg 2520
tgctgtctga tgggtaacat catgatcgac agcacgtac aacccacgcc ccagtgcacc 2580
cgcggaacg acaacctgcg gtccatgggc attggcatgc agggcctgca cacggcgtgc 2640
ctcaagatgg gcctggatct ggagtcggcc gagttccggg acctgaacac acacatcgcc 2700
gaggtgatgc tgcctcgggc catgaagacc agtaacgcgc tgtgcgttcg cggggcgcg 2760
cccttcagcc actttaagcg cagcatgtac cgggcccggc gctttcactg ggagcgcttt 2820
tcgaacgcca gcccgcggta cgaggcgag tgggagatgc tacgccagag catgatgaaa 2880
cacggcctgc gcaacagcca gttcatcgcg ctcatgcccc ccgcccctc ggcccagatc 2940
tcggacgtca gcgagggtt tgccccctg ttaccaaac tgttcagcaa ggtgaccagg 3000
gacggcgaga cgtgcgccc caacacgctc ttgtgaagg aactcgagcg cacgttcggc 3060
gggaagcggc tcctggacgc gatggacgg ctcgagggca agcagtggtc tgtggcccag 3120
gccctgcctt cctgggaccc gcgccacccc ctccggcggt tcaagacggc cttcgactac 3180
gaccaggaac tgcgtatcga cctgtgtgca gaccgcgccc cctatgttga tcacagccaa 3240
tccatgactc tgtatgtcac agagaaggcg gacgggacgc tccccgctc caccctggtc 3300
cgcttctcg tccacgcata taagcgggc ctgaagacgg ggatgtacta ctgcaaggtt 3360

```

cgcaaggcga ccaacagcgg ggtgttcgcc ggcgacgaca acatcgtctg cacaagctgc 3420
gcgctgtaa 3429

<210> 66
<211> 825
<212> DNA
<213> HSV-2

<400> 66
ggaaagtcag cggcgtgatg gtgttttcca gcatccccc cggccccgcg gcctaccgca 60
ttagcgacag cagctttgtt caatgoggct ccaactgcag tatgataatc gacggagacg 120
tggcgcgcgg tcatttgcgt gacctcgagg gcgctacgtc caccggcgcc ttctgcgcga 180
tctcaaacgt cgcagccggc ggggatggcc gaaccgcgt cgtggcgctc ggcggaacct 240
cgggcccgtc cgcgactaca tccgtgggga ccagacgtc cggggagtcc ctccacggga 300
acccaaggac cccgaacccc caaggacccc aggtgtgtcc cccgccccct cctccccctt 360
ttccatgggg ccacgagtgc tgcgccgtc gcatgccag gggcgcgcc gagaaggacg 420
tcggggccgc ggagtcattg tcagacggcc cgtcgtccga ctccgaaacg gaggactcgg 480
actcctcgga cgaggatacg ggctcggtt cggagacgct gtctcgatcc tcttcgatct 540
gggcccgcagg ggcgactgac gacgatgaca ggcactccga ctgcgggtcg gacgactccg 600
tgcagcccga cgttgtcgtt cgtcgcagat ggagcgacgg ccccgcccc gtggcctttc 660
ccaagccccg gcgccccggc gactcccccg gaaacccccg cctgggcgcc ggcaccgggc 720
cgggtccgc gacggaccgc cgcgcgtcgg ccgactccga ttccgcggcc caccgcgcgc 780
caccccaggc ggacgtggcg ccggttcttg acagccagcc cactg 825

<210> 67
<211> 678
<212> DNA
<213> HSV-2

<400> 67
atggccaacc gccctgcgc atccgccctc gccggagcgc ggtctccgtc cgaacgacag 60
gaacccccgg agcccgaggt cccccccct ggccggcacc acgtgttttg caggaaagtc 120
agcggcgtga tgggtgtttc cagcgatccc cccggccccg cggcctaccg cattagcgac 180
agcagctttg ttcaatgcgg ctccaactgc agtatgataa tcgacggaga cgtggcgcg 240
ggtcatttgc gtgacctcga ggcgcgtacg tecaccggcg ccttcgtcgc gatctcaaac 300
gtcgcagccg gcggggatgg ccgaaccgcc gtctgtggcg tcggcggaac ctggggcccg 360
tccgcgacta catccgtggg gaccagacg tccggggagt tcctccacgg gaacccaagg 420
acccccgaac cccaaggacc ccaggtgtc cccccgccc ctctccccc ctttccatgg 480
ggccacgagt gctgcgcccg tcgcgatgcc agggcgggcg ccgagaagga cgtcggggcc 540
gcggagtcat ggtcagacgg ccgctcgtcc gactccgaaa cggaggactc ggactcctcg 600
gacgaggata cgggctcggg ttccggagacg ctgtctcgat cctcttcgat ctgggcccga 660
ggggcgactg acgacgat 678

<210> 68
<211> 313
<212> DNA
<213> HSV-2

<400> 68
gacgaggggt cggaatccaa aggacgcaga ccacctttgg ttacggaccc ctttctcccc 60
cccttccgaa caaaaagcag cggcgggggg gccggggtga gggagggaca cgggggacac 120
ggcacggggg tccgcctca cgcgccgcgc cctctaaatc cccccggtt ctttgtcaag 180
cagccccgcc cccgcacgc ctgggggatg ctcaacgaca tgcagtggct cgccagcagc 240
gactcggagg aggagaccga ggtgggaatc tctgacgacg accttcaccg cgactccacc 300
tccgagggcg gca 313

<210> 69
<211> 467

<212> DNA
<213> HSV-2

<220>
<221> misc_feature
<222> 39,322,332,368,369
<223> n = A,T,C or G

<400> 69
cagggcagcc ccacgcccgc cgacgcgcag ggatcctgnt gggggtgggc ccgtgggtga 60
ggaggaagcg gaagcgggag gggggggcga cgtgtgcgcc gtgtgcacgg acgagatcgc 120
cccgcccctg cgtgccaga gttttccctg cctgcacccc ttctgcatcc cgtgcatgaa 180
gacctggatt ccgttgcgca acacgtgtcc cctgtgcaac accccgggtg cgtacctgat 240
agtgggcgtg accgccagcg ggtcgttcag caccatcccg atagtgaacg acccccgac 300
ccgcgtggag gccgaggcgg cngtgcgggt cnggcacggc cgtggacttt atctggacgg 360
gcaacccnng gacggccccg cgctccctgt cgctgggggg acacacggtc cgcgccctgt 420
cgcccacccc cccgtggccc ggcacggacg acgaggacga tgacctc 467

<210> 70
<211> 204
<212> DNA
<213> HSV-2

<220>
<221> misc_feature
<222> 78,79,120,121,124,125
<223> n = A,T,C or G

<400> 70
gcccctattg gtcccctggg cttectagta tgctaataaa tttctccccg ggggcgggca 60
ccactcaggg ccgcgcgnng ggccgcgggg gactcccacg tgcgtcggcg gggggcggn 120
natnntaatg gggttcttgg agtacacccg gttggtcccc ggggacggcc cggcccagaa 180
gggggattcc ctccctccgc cccc 204

<210> 71
<211> 474
<212> DNA
<213> HSV-2

<220>
<221> misc_feature
<222> 7,43,56,339,424,431,451,468,474
<223> n = A,T,C or G

<400> 71
ccccggnccg ctttaagcgt cgggggaccc ccgtgggccc tgngccgccc cccgancctc 60
tgggggggcg agggaggcag ggaggagccc gagagcgggg gacagggggg gagacgaggg 120
gtcggaatcc aaaggacgca gaccaccttt gggtacggac ccctttctcc ccccttccg 180
aacaaaaagc agcgggcggg gggccggggg gagggaggga cacgggggac acggcacggg 240
ggtccgcgct cagcggccgc gccctctaaa tccccccgt ttctttgtca agcagcccgc 300
cgccccgcac gcctggggga tgctcaacga catgcagtng ctgccagca gcgactcgga 360
ggaggagacc gaggtgggaa tctctgacga cgacctcac cgcgactcca cctccgaggc 420
gggncagcac nggacacgga gatgttcgag ncgggcctga tggacgcngc cacn 474

<210> 72
<211> 350
<212> DNA
<213> HSV-2

<220>

<221> misc feature

<222> 107, 148, 185, 187, 305, 312, 313

<223> n = A, T, C or G

<400> 72

```

ggggtcggaa tccaaaggac gcagaccacc tttggttacg gacccctttc tccccccctt 60
ccgaacaaaa agcagcgggc ggggggcccgg ggtgaggag ggacacnggg ggacacggca 120
cgggggtccc gcctcacgcc ccgcgccttc taaatcccc ccgtttcttt gtcaagcagc 180
ccgengnecc gcacgcctgg gggatgctca acgacatgca gtggctcgcc agcagcgact 240
cggaggagga gaccgaggtg ggaatctctg acgacgacct tcaccgcgac tccacctccg 300
aggcngggca gnnccgacac ggagatgttc gaggcgggct tgatggacgc 350

```

<210> 73

<211> 312

<212> DNA

<213> HSV-2

<220>

<221> misc feature

<222> 21, 32, 39, 66, 306

<223> n = A, T, C or G

<400> 73

```

gccccacgcc cgcgcacgcg nagggatcct gngggggtng gcccggtgggt gagggaggaag 60
cggaancggg aggggggggc gacgtgtgcg ccgtgtgcac ggacgagatc gccccgcccc 120
tgcgtgccca gagttttccc tgcctgcacc cttctgcat cccgtgcatg aagacctgga 180
ttcgttgcg caacacgtgt cccctgtgca acaccccggt ggcgtacctg atagtgggcy 240
tgaccgccag cgggtcgttc agcaccatcc cgatagtga aagaccccgg acccgcggtg 300
aggccngagg cg 312

```

<210> 74

<211> 274

<212> PRT

<213> HSV-2

<400> 74

```

Lys Val Ser Gly Val Met Val Leu Ser Ser Asp Pro Pro Gly Pro Ala
          5              10              15

```

```

Ala Tyr Arg Ile Ser Asp Ser Ser Phe Val Gln Cys Gly Ser Asn Cys
          20              25              30

```

```

Ser Met Ile Ile Asp Gly Asp Val Ala Arg Gly His Leu Arg Asp Leu
          35              40              45

```

```

Glu Gly Ala Thr Ser Thr Gly Ala Phe Val Ala Ile Ser Asn Val Ala
          50              55              60

```

```

Ala Gly Gly Asp Gly Arg Thr Ala Val Val Ala Leu Gly Gly Thr Ser
          65              70              75              80

```

```

Gly Pro Ser Ala Thr Thr Ser Val Gly Thr Gln Thr Ser Gly Glu Phe
          85              90              95

```

```

Leu His Gly Asn Pro Arg Thr Pro Glu Pro Gln Gly Pro Gln Ala Val

```

100	105	110
Pro Pro Pro Pro Pro Pro Pro Phe Pro Trp Gly His Glu Cys Cys Ala		
115	120	125
Arg Arg Asp Ala Arg Gly Gly Ala Glu Lys Asp Val Gly Ala Ala Glu		
130	135	140
Ser Trp Ser Asp Gly Pro Ser Ser Asp Ser Glu Thr Glu Asp Ser Asp		
145	150	155
Ser Ser Asp Glu Asp Thr Gly Ser Gly Ser Glu Thr Leu Ser Arg Ser		
165	170	175
Ser Ser Ile Trp Ala Ala Gly Ala Thr Asp Asp Asp Asp Ser Asp Ser		
180	185	190
Asp Ser Arg Ser Asp Asp Ser Val Gln Pro Asp Val Val Val Arg Arg		
195	200	205
Arg Trp Ser Asp Gly Pro Ala Pro Val Ala Phe Pro Lys Pro Arg Arg		
210	215	220
Pro Gly Asp Ser Pro Gly Asn Pro Gly Leu Gly Ala Gly Thr Gly Pro		
225	230	235
Gly Ser Ala Thr Asp Pro Arg Ala Ser Ala Asp Ser Asp Ser Ala Ala		
245	250	255
His Ala Ala Ala Pro Gln Ala Asp Val Ala Pro Val Leu Asp Ser Gln		
260	265	270

Pro Thr

<210> 75
 <211> 226
 <212> PRT
 <213> HSV-2

<400> 75

Met Ala Asn Arg Pro Ala Ala Ser Ala Leu Ala Gly Ala Arg Ser Pro		
5	10	15
Ser Glu Arg Gln Glu Pro Arg Glu Pro Glu Val Ala Pro Pro Gly Gly		
20	25	30
Asp His Val Phe Cys Arg Lys Val Ser Gly Val Met Val Leu Ser Ser		
35	40	45
Asp Pro Pro Gly Pro Ala Ala Tyr Arg Ile Ser Asp Ser Ser Phe Val		
50	55	60
Gln Cys Gly Ser Asn Cys Ser Met Ile Ile Asp Gly Asp Val Ala Arg		
65	70	75
Gly His Leu Arg Asp Leu Glu Gly Ala Thr Ser Thr Gly Ala Phe Val		

	85		90		95
Ala Ile Ser Asn Val Ala Ala Gly Gly Asp Gly Arg Thr Ala Val Val	100		105		110
Ala Leu Gly Gly Thr Ser Gly Pro Ser Ala Thr Thr Ser Val Gly Thr	115		120		125
Gln Thr Ser Gly Glu Phe Leu His Gly Asn Pro Arg Thr Pro Glu Pro	130		135		140
Gln Gly Pro Gln Ala Val Pro Pro Pro Pro Pro Pro Phe Pro Trp	145		150		155
Gly His Glu Cys Cys Ala Arg Arg Asp Ala Arg Gly Gly Ala Glu Lys	165		170		175
Asp Val Gly Ala Ala Glu Ser Trp Ser Asp Gly Pro Ser Ser Asp Ser	180		185		190
Glu Thr Glu Asp Ser Asp Ser Ser Asp Glu Asp Thr Gly Ser Gly Ser	195		200		205
Glu Thr Leu Ser Arg Ser Ser Ser Ile Trp Ala Ala Gly Ala Thr Asp	210		215		220

Asp Asp
225

<210> 76
<211> 4125
<212> DNA
<213> HSV-2

<400> 76
atggccgctc ctgcccgcga cccccgggt taccggtaac ccgcgcccat cctgcccacc 60
ggctccatcc tgagtacgat cgaggtggcg tcccaccgca gactctttga ttttttcgcc 120
gccgtgcgct ccgacgaaaa cagcctgtat gacgtagagt ttgacgccct gctgggggtcc 180
tactgaaca ccctgtcgct cgtgcgcttt ctggagctcg gcctgtccgt ggcgtgcgtg 240
tgcaccaagt tcccggagct ggcttacatg aacgaagggc gtgtgcagtt cgagggtccac 300
cagccccctca tcgcccgcga cggcccgcac ccgctcgagc agcccgtgca taattacatg 360
acgaaggtca tcgaccgccg ggccctgaac gccgccttca gcctggccac cgaggccatt 420
gccctgctca cgggggaggg cctggacggg acgggcatta gcctgcatcg ccagctgcgc 480
gccatccagc agctcgcgcg caacgtccag gccgtcctgg gggcgtttga gcgcggcacg 540
gccgaccaga tgctgcacgt gctgttgag aaggcgctc ccctggccct gctgttgccc 600
atgcaacgat atctcgacaa cgggcgcctg gcgaccaggg ttgcccgggc gaccctggtc 660
gccgagctga agcggagctt ttgcgacacg agcttcttcc tgggcaaggc gggccatcgc 720
cgcgaggcca tcgaggcctg gctcgtggac ctgaccacgg cgacgcagcc gtccgtggcc 780
gtgccccgcc tgacgcacgc cgacacgcgc gggcgggcgg tcgacggggg gctggtcacc 840
accgccgcca tcaaacagcg cctcctgcag tccttcttga aggtggagga caccgaggcc 900
gacgtgccgg tgacctacgg cgagatggtc ttgaacgggg ccaacctcgt cacggcgctg 960
gtgatgggca aggccgtgcg gagcctggac gacgtgggcc gccacctgct ggatatgcag 1020
gaggagcaac tcgaggcgaa ccgggagacg ctggatgaac tcgaaagcgc ccccagaca 1080
acgcgcgtgc gcgcgatct ggtggccata ggcgacaggc tggcttctct ggaggccctg 1140
gagagacgca tctacgccgc caccaacgtg ccctaccccc tggtgggcgc catggacctg 1200
acgttcgtcc tgcccctggg gctgttcaac ccggccatgg agcgcttcgc cgcgcacgcc 1260
ggggacctgg tgcccggccc cggccacccg gagccccgcg cgttccctcc ccggcagctg 1320
tttttttggg gaaaggacca ccaggttctg cggctgtcca tggagaacgc ggtcgggacc 1380

gtgtgtcatc	cttcgctcat	gaacatogac	gcggecgteg	ggggcggtgaa	ccaacgacccc	1440
gtcgaggccg	cgaatccgta	cgggcggtac	gtcgcgggccc	cgcccgggccc	cgggcggggac	1500
atgcagcagc	gtttttctgaa	cgccctggcgg	cagcgccctcg	cccacggccg	ggtccgggtgg	1560
gtcgccgagt	gccagatgac	cgcgaggacg	ttcatgcagc	ccgacaacgc	caacctgggt	1620
ctggagctgc	accccgcggt	cgactttctt	gcgggcggtg	ccgacgtcga	gcttcccggc	1680
ggcgaagtcc	ccccggccgg	tccggggggc	atccaggcca	cctggcgcg	ggtcaacggc	1740
aacctgcccc	tggcgctgtg	tccgggtggc	tttcgtgacg	cccggggcct	ggagctcggc	1800
gttggccgcc	acgccatggc	gccggctacc	atagccgcgc	tcccgggggc	gttcgaggac	1860
cgcagctacc	cggcggtgtt	ttacctgctg	caagccgcga	ttcacggcaa	cgagcacgtg	1920
ttctgcgccc	tggcgcggt	cgtgactcag	tgcatcacca	gtactggaa	caacacgcga	1980
tgcgggcggt	tcgtgaacga	ctactcgctg	gtctcgtaca	tcgtgacct	cctcgggggc	2040
gacctcccc	aggagtgc	ggcgtgtat	cgggacgtg	tggccacgt	cgaggccctg	2100
gcccagctgg	tggacgactt	taccctggcg	ggcccgagc	tggcggggca	ggctcaggcc	2160
gagctgaatc	acctgatgcg	cgaccggcg	ctgctgccc	ccctcggtg	ggactgcgac	2220
ggccttatgc	gacacgcggc	cctggaccgc	caccgagact	gccgattga	cgcggggggg	2280
caagagcccc	tctacgcggc	ggcgtgcaac	gtggcgacgg	ccgacttta	ccgcaacgac	2340
ggccggctgc	tgcaaacac	ccaggcccgc	gcgccgacg	ccgccgacga	ccggccgcac	2400
cgccggcgcc	actggaccgt	ccaccacaaa	atctactatt	acgtgctggt	gccggccttc	2460
tcgcgggggc	gctgctgcac	cgccgggggt	cgcttcgacc	gcgtgtacgc	cacgctgcag	2520
aacatggtgg	tcccggagat	cgcccccggt	gaggagtgc	cgagcgatcc	cgtgaccgac	2580
cccgcgccacc	cgtgcatcc	cgccaatctg	gtggccaaca	cgtcaagcg	catgttcac	2640
aacgggcgcg	tcgtcgtcga	cgggcccgc	atgctcacgc	tgcaggtgct	ggcgcaaac	2700
atggccgagc	gcacgacggc	gctgctgtgc	tccgcggcgc	ccgacgcggg	cgccaacacc	2760
gcgtcgacgg	ccaacatg	catcttcgac	ggggcgctgc	acgccggcgt	gctgctcatg	2820
gccccccagc	acctggacca	caccatccaa	aatggcgaat	acttctacgt	cctgcccgtc	2880
cacgcgctgt	ttgcgggcgc	cgaccacggt	gccaacgcgc	ccaacttccc	cccggccctg	2940
cgcgacctgg	cgcgcgacgt	ccccctggtc	cccccgggcc	tgggggcca	ctactttctc	3000
tccatccgcc	agcccggtg	gcagcacgcc	cgcgagagcg	cgccggggga	gaacgcgctg	3060
acctacgcgc	tcattggcgg	gtacttcaag	atgagccccg	tggccctgta	tcaccagctc	3120
aagaogggcc	tccaccccgg	gttcgggttc	accgtcgtgc	ggcaggaccg	cttcgtgacc	3180
gagaacgtgc	tgttttccga	gcgcgcgctg	gaggcgact	ttctgggcca	gctccagggtg	3240
geccgccacg	aaacgggcgg	gggggtcaac	ttcacgctca	cccagcccg	cggaacgtg	3300
gacctgggtg	tgggtacac	cgccgtcgcg	gccacgggca	ccgtccgcaa	ccccgttacg	3360
gacatgggca	acctccccca	aaacttttac	cctggccgcg	ggggcccccc	gctgctagac	3420
aacgcggccg	ccgtgtacct	gcgcaacgcg	gtcgtggcgg	gaaaccggct	ggggccggcc	3480
cagccctccc	cggctcttgg	ctgcgcccag	gtgcgcggc	gcgcggcat	ggaccaacggg	3540
caggatggcg	tgtgtgagtt	catcgccacc	cccggtggca	cggacatcaa	ctactttcgc	3600
cgccctgca	accccggggg	acgcgcggcc	ggcgcggtgt	acgcggggga	caaggagggg	3660
gagctcatag	ccctcatgta	cgaccaacggc	cagagcgacc	cgccgcggcc	cttcgcggcc	3720
acggccaacc	cgtgggcgtc	gcagcgggtt	tcgtacgggg	acctgctgta	caacggggcc	3780
tatcacctca	acggggcctc	gcccgtcctc	agcccctgct	tcaagttctt	cacgcgggcc	3840
gacatcacgg	ccaaacatcg	ctgcctggag	cgtctcatcg	tggaaacggg	atcgcggtg	3900
tccacggcca	ccgctgccag	cgacgtgcag	tttaagcgcc	cgccgggggtg	ccgcgagctc	3960
gtggaagacc	cgtgcggcct	gtttcaggaa	gcctaccgca	tcacctgcgc	cagcgacccc	4020
gccctgctac	gcagcgccc	cgatggggg	gcccacgcgc	gagagaccca	ctttacgcag	4080
tatctcatct	acgacgcctc	cccgtctaa	ggcctgtctc	tgtaa		4125

<210> 77

<211> 8108

<212> DNA

<213> HSV-2

<400> 77

ttcttgaaga	cgaaagggcc	tcgtgatacg	cctattttta	taggttaatg	tcatgataat	60
aatggttttc	tagacgtcag	gtggcacttt	tcggggaaat	gtgcgcggaa	cccctatttg	120
ttttattttc	taaatacat	caaataatgta	tccgctcatg	agacaataac	cctgataaat	180
gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttcgctg	tcgcccttat	240
tcoccttttt	gcggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	tggtgaaagt	300

aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	atctcaacag	360
cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	gcacttttaa	420
agttctgcta	tgtggcgcg	tattatcccg	tattgacgcc	gggcaagagc	aactcggctc	480
ccgcatacac	tattctcaga	atgacttgg	tgagtactca	ccagtcacag	aaaagcatct	540
tacggatggc	atgacagtaa	gagaattatg	cagtgcctgc	ataacatga	gtgataacac	600
tgcggccaac	ttacttctga	caacgatcgg	aggaccgaag	gagctaaccg	cttttttgca	660
caacatgggg	gatcatgtaa	ctcgccctga	tcgttgggaa	cggagctga	atgaagccat	720
accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	gcaacaacgt	tgcgcaact	780
attaactggc	gaactactta	ctctagcttc	cgggcaacaa	ttaatagact	ggatggaggc	840
ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttcgg	gctggctggt	ttattgctga	900
taaactctga	gccggtgagc	gtgggtctcg	cggtatcatt	gcagcactgg	ggccagatgg	960
taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	caggcaacta	tggatgaacg	1020
aaatagacag	atcgctgaga	taggtgcctc	actgattaag	cattggtaac	tgtcagacca	1080
agtttactca	tataactttt	agattgattt	aaaacttcat	ttttaattta	aaaggatcta	1140
ggtgaagatc	ctttttgata	atctcatgac	caaaatccct	taacgtgagt	tttcgttcca	1200
ctgagcgtca	gaccccgtag	aaaagatcaa	aggatcttct	tgagatcctt	ttttctgcg	1260
cgtaatctgc	tgcttgcaaa	caaaaaaacc	accgctacca	gcggtggttt	gtttgccgga	1320
tcaagagcta	ccaactcttt	ttccgaaggt	aactggcttc	agcagagcgc	agataccaaa	1380
tactgtcctt	ctagtgtagc	cgtagttagg	ccaccacttc	aagaactctg	tagcaccgcc	1440
tacatacctc	gctctgctaa	tcctgttacc	agtggctgct	gccagtggcg	ataagtcgtg	1500
tcttacccgg	ttggactcaa	gacgatagtt	accggataag	gcgcagcgtt	cgggctgaac	1560
ggggggttcg	tgcacacagc	ccagcttgga	gcgaacgac	tacaccgaac	tgagatacct	1620
acagcgtgag	cattgagaaa	gcgccacgct	tcocgaaggg	agaaaggcgg	acaggtatcc	1680
ggtaagcggc	agggtcgga	caggagagcg	cacgagggag	cttcaggggg	gaaacgcctg	1740
gtatctttat	agtccgtgct	ggtttcgcca	cctctgactt	gagcgtcgat	ttttgtgatg	1800
ctcgtcaggg	gggaggagcc	tatggaaaaa	cgccagcaac	gcggcctttt	tacggttcct	1860
ggccttttgc	tggccttttg	ctcacatgtt	ctttcctgcg	ttatcccctg	attctgtgga	1920
taaccgtatt	accgcctttg	agtgcgctga	taccgctcgc	cgagccgaa	cgaccgagcg	1980
cagcgagtca	gtgagcgagg	aagcgggaag	gcgcccaata	cgcaaaccgc	ctctccccgc	2040
gcgttggccg	attcattaat	gcagctggca	cgacagggtt	cccgaactga	aagcgggcag	2100
tgagcgcaac	gcaattaatg	tgagttagct	cactcattag	gcaccccagg	ctttacactt	2160
tatgcttcgg	gctcgtatgt	tgtgtggaat	tgtgagcgga	taacaatttc	acacaggaaa	2220
cagctatgac	catgattacg	ccaagctttt	gcgatcaata	aatggatcac	aaccagtatc	2280
tcttaacgat	gttcttcgca	gatgatgatt	catttttttaa	gtatttggct	agtcaagatg	2340
atgaaatctt	cattatctga	tatattgcaa	atcactcaat	atctagactt	tctgttatta	2400
ttattgatcc	aatcaaaaaa	taaattagaa	gocgtgggtc	attgttatga	atctctttca	2460
gaggaatata	gacaattgac	aaaattcaca	gactttcaag	attttaaaaa	actgtttaac	2520
aaggctcccta	ttgttacaga	tgggaagggtc	aaacttaata	aaggatattt	gttcgacttt	2580
gtgattagtt	tgatgcgatt	caaaaaagaa	tcctctctag	ctaccaccgc	aatagatcct	2640
gttagataca	tagatcctcg	tcgcaatatc	gcatttttcta	acgtgatgga	tatattaaag	2700
tcgaataaag	tgaacaataa	ttaattcttt	attgtcatca	tgaacggcgg	acatattcag	2760
ttgataatcg	gccccatggt	ttcaggtaaa	agtacagaat	taattagacg	agttagacgt	2820
tatcaaatag	ctcaatataa	atgcgtgact	ataaaatatt	ctaacgataa	tagatacggg	2880
acgggactat	gagcgcattg	taagaataat	tttgaagcat	tggaaagcaac	taaaactatgt	2940
gatctcttgg	aatcaattac	agatttctcc	gtgataggta	tcgatgaagg	acagttcttt	3000
ccagacattg	ttgaattccg	agcttggctg	caggctgggg	atccccctg	cccgtttatt	3060
attatttttg	acaccagacc	aactggtaat	ggtagcgacc	ggcgctcagc	tgggaattccg	3120
ccgatactga	cgggctccag	gagtcgtcgc	caccaatccc	catatggaaa	ccgtcgatat	3180
tcagccatgt	gccttcttcc	gcgtgcagca	gatggcgatg	gctggtttcc	atcagttgct	3240
gttgactgta	gcggctgatg	ttgaactgga	agtcgccgcg	ccactggtgt	gggccataat	3300
tcaattcgcg	cgtcccgcag	cgagaccggt	tttcgctcgg	gaagacgtac	ggggtataca	3360
tgtctgacaa	tggcagatcc	cagcgggtcaa	aacaggcggc	agtaaggcgg	tcgggatagt	3420
tttcttgccg	ccctaattccg	agccagttta	cccgtctcgc	tacctgcgcc	agctggcagt	3480
tcaggccaat	ccgcgcggga	tgcggtgtat	cgctgcgcac	ttcaacatca	acggtaatcg	3540
ccatttgacc	actaccatca	atccggtagc	tttccggct	gataaataag	gttttccctt	3600
cgatttgcca	cgcgtgagcg	gtcgtaatat	gcaccgcatc	agcaagtgtg	tctgcgctgc	3660
actgcaacaa	cgtctcttcg	gcctggtaat	ggccgcgcgc	cttcagcgt	tcgaccaggg	3720
cgttagggtc	aatgcgggtc	gcttcaactta	cgccaatgtc	gttatccagc	ggtgcacggg	3780

tgaactgata	gdcagcggc	gtcagcagtt	gttttttata	gccaatccac	atctgtgaaa	3840
gaaagcctga	ctggcgggta	aattgccaac	gcttattacc	cagctcgatg	caaaaatcca	3900
tttcgctggg	ggtcagatgc	gggatggcgt	gggacgcggc	ggggagcgtc	acactgaggt	3960
tttcgccag	acgccactgc	tgccaggcgc	tgatgtgccc	ggcttctgac	catgagggtc	4020
cggtcgggtg	cactacgcgt	actgtgagcc	agagttgccc	ggcgtctctc	ggctgaggta	4080
gttcaggcag	ttcaatcaac	tgtttaccct	gtggagcgac	atccagaggc	acttcaccgc	4140
ttgccagcgg	cttaccatcc	agcgccacca	tccagtgcag	gagctcgta	tcgctatgac	4200
ggaacaggta	ttcgctggtc	acttcgatgg	tttgcccggg	taaacgggaa	tggaaaaact	4260
gctgctgggt	ttttgcttcc	gtcagcgctg	gatgcggcgt	gcggtcggca	aagaccagac	4320
cggtcatata	gaactggcga	tcggtcggcg	tatcgccaaa	atcaccgccc	taagccgacc	4380
acgggttgcc	gttttcatca	taattaatca	gcgactgata	caccagtgcc	cagacgaagc	4440
cgccctgtaa	acggggatac	tgacgaaacg	cctgccagta	tttagcgaaa	ccgccaagac	4500
tgttaccat	cgcgtgggcg	tattcgcaaa	ggatcagcgg	gcgctctctc	ccaggtagcg	4560
aaagccattt	tttgatggac	catttcggca	cagccgggaa	gggctgggtc	tcattccacg	4620
gcgcgtacat	cgggcaata	atatcggtgg	ccgtgggtgc	ggctccgccc	ccttcatact	4680
gcaccgggcg	ggaaggatcg	acagatttga	tccagcgata	cagcgcgtcg	tgattagcgc	4740
cgtggcctga	ttcattcccc	agcgaccaga	tgatcacact	cgggtgatta	cgatcgcgct	4800
gcaccattcg	cgttacgcgt	tcgctcatcg	ccggtagcca	gcgcggatca	tcggtcagac	4860
gattcattgg	caccatgccg	tggttttcaa	tattggcttc	atccaccaca	tacaggccgt	4920
agcggtcgca	cagcgtgtac	cacagcggat	ggttcggata	atgcgaacag	cgcacggcgt	4980
taaagtgtgt	ctgcttcata	agcaggatat	cctgcacat	cgtctgctca	tccatgacct	5040
gaccatggcag	aggatgatgc	tcgtgacggg	taacgcctcg	aatcagcaac	ggcttgccgt	5100
tcagcagcag	cagaccattt	tcaatccgca	cctcgcgga	accgacatcg	caggcttctg	5160
cttcaatcag	cgtgccgtcg	gcgggtgtga	gttcaaccac	cgcacgatag	agattcggga	5220
tttcggcgct	ccacagtttc	gggttttcga	cggtgagacg	tagtgtgacg	cgatcgccat	5280
aaccaccacg	ctcatcgata	atttcaccgc	cgaaaggcgc	gggtccgctg	gcgacctgcg	5340
tttcaccctg	ccataaagaa	actgttacct	gtaggtagtc	acgcaactcg	ccgcacatct	5400
gaacttcagc	ctccagtaca	gcgcggctga	aatcatcatt	aaagcagtg	gcaacatgga	5460
aatcgctgat	ttgtgtagtc	ggtttatgca	gcaacgagac	gtcacggaaa	atgccgctca	5520
tcggccacat	atcctgatct	tccagataac	tgccgtcact	ccaacgcagc	accatcaccg	5580
cgaggcgggt	ttctccggcg	cgtaaaaatg	cgcctcaggc	aaattcagac	ggcaaacgac	5640
tgctctggcc	gtaaccgacc	cagcgcccg	tgaccacag	atgaaacgcc	gagttaacgc	5700
catcaaaaat	aattcgcgtc	tgcccttct	gtagccagct	ttcatcaaca	ttaaatgtga	5760
gcgagtaaca	accgctcgga	ttctccgtgg	gaacaaacgg	cggattgacc	gtaatgggat	5820
aggttacgtt	ggtgtagatg	ggcgcatcgt	aaccgtgcac	ctgccagttt	gaggggacga	5880
cgacagtata	ggcctcagga	agatcgcact	ccagccagct	ttccggcacc	gcttctgggt	5940
ccggaaacca	ggcaaaagcg	cattcgccat	tcaggctgcg	caactgttgg	gaagggcgat	6000
cggtgcgggc	ctcttcgcta	ttacgccagc	tgccgaaagg	gggatgtgct	gcaaggcgat	6060
taagttgggt	aacgccaggg	ttttccaggt	cacgacgttg	taaaacgacg	ggatccctcg	6120
aggaattcat	ttatagcata	gaaaaaaaca	aatgaaatt	ctactatatt	tttcatata	6180
tatatcttaa	atatgaaagt	ggtgattgtg	actagcgtag	catcgcttct	agacatata	6240
tatatagtaa	taccaatact	caagactacg	aaactgatac	aatctcttat	catgtgggta	6300
atgttctcga	tgtcgaatag	ccatatgccg	gtagttgcga	tatacataaa	ctgatcacta	6360
attccaaaacc	caaccgcttt	ttatagtaag	tttttcccc	ataaataata	aatacaataa	6420
ttaatttctc	gtaaaagtag	aaaatatatt	ctaatttatt	gcacggtaag	gaagtagaat	6480
cataaagaac	agtgcaggat	cccgtaaaac	gacggccagt	gagcgcgct	aatacgactc	6540
actatagggc	gaattgggta	ccgggcccc	cctcgagggt	gacgggtatc	ataagcttga	6600
tatcgaattc	ctgcagcccg	ggggatccac	tagttctaga	gcggccgcca	ccgagggtga	6660
gctccagctt	ttgttccctt	tagtgagggt	taattgcgcg	cttggcgtaa	tcattggtcat	6720
agctgttggg	aattctgtga	gcgtatggca	aacgaaggaa	aaattagtta	tagtagccgc	6780
actcgatggg	acatttcaac	gtaaaccggt	taataatatt	ttgaatctta	ttccattatc	6840
tgaaatggtg	gtaaaactaa	ctgctgtgtg	tatgaaatgc	tttaaggagg	cttccttttc	6900
taaacgattg	ggtgaggaaa	ccgagataga	aataatagga	ggtaatgata	tgatcaaatc	6960
ataaacattg	aagtgttaca	tcgactcata	atattatatt	ttttatctaa	aaaactaaaa	7020
ccgtttatgt	attttgagga	aatataatac	ttaaaaatgg	atgttgtgtc	gttagataaa	7080
gaggtcgcaa	aaaaactgcc	gtatcaagga	gagttagatt	acgaaccaga	aagtgcacaa	7140
tttcttagta	agttacagcg	acacgggtata	cagttaaaac	tattactagg	agaattatct	7200
			ttagatgggtg	ccaccgtagt	gtatatagga	7260

```

tctgctcccg gtacacatat acgttatttg agagatcatt tctataattt aggagtgate 7320
atcaaatgga tgctaattga cggccgccat catgatccta ttttaaattg attgcgtgat 7380
gtgactctag tgactcgggt cgttgatgag gaatatctac gatccatcaa aaaacaactg 7440
catccttcta agattatttt aatttctgat gtgagatcca aacgaggagg aaatgaacct 7500
agtacggcgg atttactaag taattacgct ctacaaaatg tcatgattag tatttttaaac 7560
cccgtggcgt ctagtcttaa atggagatgc ccgtttccag atcaatggat caaggacttt 7620
tatatcccac acggttaataa aatgttacia ccttttgctc cttcatattc agggccgctg 7680
ttttacaacg tcgtgactgg gaaaaccctg gcgttaccca acttaatcgc cttgcagcac 7740
atcccccttt cgcagctgg cgtaatagcg aagaggcccg caccgatcgc cttcccaac 7800
agttgcgcag ctgaatggc gaatggcgc tgatgcggta ttttctcttt acgcatctgt 7860
gcggtatttc acaccgata tggcgactc tcagaccat ctgctctgat gccgcatagt 7920
taagccagta cactccgcta tcgctacgtg actgggtcat ggctgcgccc cgacaccgcg 7980
caacaccgcg tgacgcgccc tgacgggctt gtctgctccc ggcatccgct tacagacaag 8040
ctgtgaccgt ctccgggagc tgcattgtgc agaggttttc accgtcatca ccgaaacgcg 8100
cgaggcag

```

<210> 78

<211> 2091

<212> DNA

<213> HSV-2

<400> 78

```

atgtccgtgc gcgggcatgc cgtacgccgg aggcgcgcct ccaccgcgtc ccatgccccg 60
tccgcgcacg ggcgcgactc gcccggtggg gacgagcccg agggcggtgg agtcgggtta 120
atgggggtacc tgcgggcgggt gttaaactg gacgacgaca gcgaggtcga ggccgcgggg 180
gagatggcga gcgaagagcc gccccgcgc cgtgcgcggg agggccgcgg tcaccccggg 240
tccgcacgcg cgtccgaggg ccgggcggcg gcgccccccc gccgggcgtc ctttccgcgc 300
cccaggtccg ttacggccag gagccagtcc gttcgcggac gccgggacag cgccatcacg 360
cgggccccgc ggggaggcta cctgggcccg atggaccac gcgacgtttt ggggcgggtg 420
ggcggttcgc ggggtggtgc ctcgcgcgtg ttcttgagc agctcaacta cgaggaggac 480
gactaccctg ccgcgcgcgc gcacgatgac ggccccggg cgcgcccttc cgcgacggtc 540
gagattctcg cgcgcgcggt gtgcggcccc gagctgcagg cggcattccc cctggaccgc 600
ctgacccccg gagtgcgcgc gtgggacgag tccgtgcgtt cggccctggc cctgggacat 660
ccggccgggt tctaccctg tccggatagc gcgttcgggc tgtcgcgcgt ggggggtcatg 720
cactttgcct ccccgccga cccaaagggt ttttccgcc agacgctgca gcagggcgag 780
gcgctggcct ggtacgtcac gggcgacgcg atcctcgacc tgacggatcg gcgggcaaaa 840
accagccct cccgcgcgat gggttttctg gtggacgcca tcgtgcgggt ggcgatcaac 900
gggtgggtct gcgggacgcg cctgcacacg gaggggcgcg gctcgagct cgacgacagg 960
gcggccgagc tccgacggca gttcgcgagc ctacggcgt tgcggcccgt gggggccgcg 1020
gccgtgccgc tgcctcagcg gggaggggcc gcgccccccc accccggccc cgacgcgcgcg 1080
gtctttcgca gttcgttgg gtccctgctg tactggccc ggtgcgcgc gctcctgggg 1140
cgcgactgtc gcgtggccgc ccgctacgcg gggcgcatga cgtacatcgc caccggggct 1200
ctgctcgccc gcttcaaccc cggcgccgtc aaatgcgtgc tcccgcggga ggccgcgtt 1260
gcggggcgcg tccgtgacgt gctggcggtc ctggcgagc agacggtcca gtggctctcg 1320
gtggctcgtg gggcgcgcc gcacccgcac tccgcccacc ccgcgtttgc ggacgtggag 1380
caggaggcgc tgtttcgcg cctgcccctg ggcagcccc gggtcgtggc ggccgagcac 1440
gaggcgctgg gcgacaccgc ggcgcgcgc ctgctcgcca ccagcgggct gaacgcctg 1500
ctgggcgcgg ccgtgtacgc gctgcacacg gccctggcga ccgttaccct gaaatacgcc 1560
ctggcctgcg gggacgcgcg ccggcgagc gacgacgcg cgccgcgcgc cgcgtgctg 1620
gcgacggggc tcatcctgca gcggtgctg ggectggcc acacgggtgg cgcgtgctg 1680
gccctggccg cgtttgacgg cgggtcgacg gccccgagg tgggcacgta cccccctg 1740
cgctacgcgt gcgtcctcc cgcgacccag cccctgtac cgcggaccac ccccgccaaa 1800
ttttgggcgc acgtgcgcgc cgccgcggaa cacgtggacc ttcgccccgc gtcctcgggc 1860
ccccggcgcc ccgtgcgcg gacggcagac cccgccttc tgcgcaaga cctggcgccc 1920
ttccccccg cccccctgaa tagcgagtcc gtgctgggg cgcgggccg cgtcgtggac 1980
atcatggcgc agtttcggaa actgctcatg ggcgacgagg agaccgcgc cctccggggc 2040

```


cacgtgtccg ggaggcgcgc gaccgggctg ggcgcccgcc cagccccata g 2091

<210> 79
<211> 1110
<212> DNA
<213> HSV-2

<400> 79
atgagtcagt gggggcccg ggcgacccct gtccagacgg acagcaccaa ccggaatgcc 60
gatggggact ggcaagcggc cgtagctatt cgcgggggag gagtcgttca actgaacatg 120
gtcaacaaac gcgcgctgga ttttaccctg gcagaatgcg gggactccga atgggcccgtg 180
ggccgcgtct ctctgggcct gcgaatggca atgcgcggg acttctgcgc gattattcac 240
gcccccgagg tatccggccc cgggccccac gtgatgctcg gtctcgtcga ctccgggctac 300
cgcggaaccg tccctggccgt ggtcgtagcc ccgaacggga cgcgcgggtt tgcccccggg 360
gccctccggg tcgacgtgac gtttctggac atccggggca ccccccgac cctcaccgag 420
ccgagctccc tgcaccggtt tccgcagttg gcgcgctccc cgctggcagg gttacgagaa 480
gatccttggt tggacggggc gctcgcgacc gccggggggg cgggtggccct gccggccaga 540
cggcgcgggg gatcgcctgg ctacgcgggc gagctaaccg aggtgaccac cgagcacggc 600
gactgcgtgc acgagggccc cgcctttctg ccaaagcgcg aggaggacgc aggttttgac 660
attctcatcc accgagccgt gaccgtcccg gccaacggcg ccacggtcac acagccgtcc 720
ctccgcgtat tgcgcgcggc cgacggacca gaggcctgct atgtgctggg gcggtcgtcg 780
ctcaatgccca ggggcctcct ggtcatgcct acgcgctggc cctccgggca cgcctgtgcg 840
tttgtgtat gtaacctgac cggagtcccgt gtgacctac aagccgggtc caaggtcgcc 900
cagctgctcg tcgcggggac ccacgccctc ccttgatcc cccccgaaa catccacgag 960
gacggcgcac tccggggcta cccagaggg gtcccgagac cgaccgccac cccccgagac 1020
ccgcccattt tgggtgtttac gaacgagttt gacgcggacg ccccccaaag caagcggggg 1080
gccggggggg ttggctccac tggcatctag 1110

<210> 80
<211> 228
<212> DNA
<213> Homo sapiens-ubiquitin UL49-HSV-2

<400> 80
atgcagatct tcgtgaagac tctgactggg aagaccatca ccctcgaggg ggagcccagt 60
gacaccatcg agaattgtcaa ggcaaagatc caagataagg aaggcattcc tcctgatcag 120
cagaggttga tctttgcccg aaaacagctg gaagatgggc gtaccctgtc tgactacaac 180
atccagaaaag agtccacctt gcacctggta ctccgtctca gaggtggg 228

<210> 81
<211> 903
<212> DNA
<213> Homo sapiens-ubiquitin UL49-HSV-2

<400> 81
atgacctctc gccgctccgt caagtcgtgt ccgcgggaag cgcgcgcgag gaccacagag 60
gagctgtact atggcccggg ctccccggcg gaccagaga gtccgcgcga cgacttcgcg 120
cgcggcgcgt gcccgatgcg cgcgcgcccg aggggcgagg ttcgctttct ccattatgac 180
gaggttgggt atgccctcta ccgggactcg tcttcggacg acgacgagtc ccgggatacc 240
gcgcgaccgc gtcgttcggc gtccgtcgcg ggctctcacg gccccggccc cgcgcgcgct 300
cctccacccc ccggggggcc cgtgggcgac ggcgggcgct cgacgcgccc tcccgcgcgc 360
acccccaaaa tgacgcgcg ggccgctaag gcctccgca ccccgcgac cgaccccgcc 420
cgcggcaggc gaccgcgcca ggcgactcc cccgtgctcc tagacgccc cgctccacg 480
gcctcgggaa gaaccaagac acccgcccag ggactggcca agaagctgca cttcagcacc 540
gccccaccga gccccacggc gccgtggacc ccccggttg cggggttcaa caagcgcgctc 600

```

ttctgcgcgcg cggtcggggcg cctggcgggcc acgcacgccc ggctggcggc ggtacagctg 660
tgggacatgt cgcggccgca caccgacgaa gacctcaacg agctcctcga cctcaccacc 720
attcgcgtga cggctctgcga gggcaagaac ctctcgagc gcgcgaacga gttggtgaat 780
cccgcgcggc cgcaggacgt cgacgcgacc gcggcgcccc ggggccccgc cgcggggctg 840
gccgccgcga ccgcacgggc ccccgccccgc tcggcttccc gtccccgcgc cccctcagag 900
tag                                         903

```

<210> 82

<211> 1113

<212> DNA

<213> Homo sapiens

<400> 82

```

atgagtcagt gggggccccag ggcgatacctt gtccagacgg acagcaccaa ccggaatgcc 60
gatggggact ggcaagcggc cgtagctatt cgcgggggcg gaggcgttca actgaacatg 120
gtcaacaaac gcgcgctgga ttttaccctg gcagaatgcg gggactccga atgggcccgtg 180
ggccgcgtct ctctgggccc gcgaatggca atgcgcggcg acttctgcgc gattattcac 240
gccccgcggc tatccggccc cgggccccac gtgatgctcg gtctcgtcga ctcgggctac 300
cgcggaaccg tcctggccgt ggtcgtagcc ccgaacggga cgcgcgggtt tgccccggg 360
gccctccggg tcgacgtgac gtttctggac atccgggcca ccccccgac cctcaccgag 420
ccgagctccc tgcaccggtt tccgcagttg gcgcgctccc cgetggcagg gttacgagaa 480
gatccttggg tggacggggc gctcgcgacc gccggggggg cggggtgccc gccggccaga 540
cggcgcgggg gatcgctggg ctacgcgggc gaggtaacgc aggtgaccac cgagcacggc 600
gactgcgtgc acgaggcgcc cgcctttctg ccaaagcgcg aggaggacgc aggtttgac 660
attctcatcc accgagccgt gaccgtcccg gccaacggcg ccacggctcat acagccgtcc 720
ctccgcgtat tgcgcgcggc cgacggacca gaggcctgct atgtgctggg gcggtcgtcg 780
ctcaatgccg ggggcctcct ggtcatgcct acgcgctggc cctccgggca cgcctgtgcg 840
tttggtgtat gtaacctgac cggagtcctc gtgaccctac aagccgggtc caaggtcgcc 900
cagctgctcg tcgcggggac ccacgccttc cctggatcc cccccgacaa catccacgag 960
gacggcgcat tccgggccta cccagaggg gttccggacg cgaccgccac cccccgagac 1020
ccgccgattt tgggtgtttac gaacgagttt gacgcggacg ccccccaag caagcggggg 1080
gccggggggt ttggctccac tggcatctag tga                                         1113

```

<210> 83

<211> 927

<212> DNA

<213> Homo sapiens

<400> 83

```

atgcagcatc accaccatca ccaccacacc tctcgccgct ccgtcaagtc gtgtccgcgg 60
gaagcgccgc gcgggaccca cgaggagctg tactatggcc cggctctccc ggcggaccca 120
gagagtcgcg gcgacgactt ccgcgcgggc gctggccccg tgcgcgcgcg cccgaggggc 180
gaggttcgct ttctccatta tgacgaggct gggatatgcc tctaccggga ctctcttcg 240
gacgacgacg agtcccgggg taccgcgcca ccgcgctcgt cggcgtccgt cgcgggctct 300
cacggccccg gccccgcgog cgtctctcca cccccggggg gccccgtggg cgcgggctgg 360
cgctcgacgc cccctccgcg gcggaccccc aaaatgacgc gcggggcgcc taaggcccc 420
gcgacccccg cgaccgaacc cgcgcgcggc aggcgacccg cccaggccga ctccgcctg 480
ctcctagacg cccccgctcc cacggcctcg ggaagaacca agacacccgc ccagggactg 540
gccaaagaag tgcacttcag caccgccccg ccgagccccg cggcgccgtg gaccccccg 600
gtggccgggt tcaacaagcg cgtcttctgc gccgcggctg ggcgcctggc ggccacgcac 660
gcccggtgg cggcgggtaca gctgtgggac atgtcgcggc cgcacaccga cgaagacctc 720
aacgagctcc tcgacctcac caccattcgc gtgacggctc gcgagggcaa gaacctctg 780
cagcgcgcca acgagttggt gaatcccagc gcggcgagc acgtcgacgc gaccgcggcc 840
gcccgggggc gccccgcggg gcgtgcgcgc gcgaccgcac gggccccgcg ccgctcggt 900
tcccgctccc gccgccccct cgagtag                                         927

```

<210> 84

<211> 4149
 <212> DNA
 <213> Homo sapiens

<400> 84

```

atgcagcatc accaccatca ccacgccgct cctgcccgog accccccggg ttaccggtag 60
gccgcggcca tggtagccac cggctccatc ctgagtagca tcgaggtggc gtcccaccgc 120
agactctttg attttttcgc ccgcgtgcgc tccgacgaaa acagcctgta tgacgtagag 180
tttgacgccc tgcctggggtc ctactgcaac accctgtcgc tcgtgcgctt tctggagctc 240
ggcctgtccg tggcgtgcgt gtgcaccaag ttcccgagc tggcttacat gaacgaaggg 300
cgtgtgcagt tcgaggtcca ccagcccctc atcgcccggg acggcccgca ccccgctcag 360
cagcccgtgc ataattacat gacgaaggct atcgaccgcc gggccctgaa cgccgccttc 420
agcctggcca ccgaggccat tgccctgctc acgggggagg ccctggacgg gacgggcatt 480
agcctgcatc gccagctgcg cggcatccag cagctcgcgc gcaacgtcca ggccgtcctg 540
ggggcgtttg agcgcggcac ggccgaccag atgctgcacg tgcctgttga gaaggcgctt 600
cccctggccc tgcctgttgc catgcaacga tatctcgaca acgggcgcct ggcgaccagg 660
gtcgcccggg cgaccctggg cgccgagctg aagcggagct tttgcgacac gagcttcttc 720
ctgggcaagg cgggccatcg ccgcgaggcc atcgaggcct ggctcgtgga cctgaccacg 780
gcgacgcagc cgtccgtggc cgtgccccgc ctgacgcacg ccgacacgcg cgggcggccg 840
gtcgacgggg tgcctgtcac caccgcgcgc atcaaacagc gcctcctgca gtccttcctg 900
aaggtggagg acaccgaggc cgacgtgccc gtgacctacg gcgagatggt cttgaacggg 960
gccaacctcg tcacggcgct ggtgatgggc aaggccgtgc ggagcctgga cgacgtgggc 1020
cgccacctgc tggagatgca ggaggagcaa ctcgaggcga accgggagac gctggatgaa 1080
ctcgaaagcg cccccagac aacgcgcgtg cgcgcggatc tggtagccat aggcgacagg 1140
ctggtcttcc tggaggccct ggagaagcgc atctacgcg ccaccaacgt gccctacccc 1200
ctggtggggc ccatggacct gacgttcgtc ctgcccctgg ggctgttcaa cccggccatg 1260
gagcgttcg ccgcgcacgc cggggacctg gtgcccgcgc ccggccaccc ggagccccgc 1320
gcgttccctc cccggcagct gtttttttgg ggaaggacc accaggttct gcggctgtcc 1380
atggagaacg cggtcggggc cgtgtgtcat ccttcgtca tgaacatcga cgcgccgctc 1440
gggggcgtga accacgcccc cgtcgaggcc gcgaaccctg acggggcgta cgtcgcgccc 1500
ccggccggcc ccggcgcgga catgcagcag cgttttctga acgcctggcg gcagcgctc 1560
gccacggccc gggtcgggtg ggtcgccgag tgcagatga ccgcgagca gttcatgcag 1620
cccgaacag ccaacctggc tctggagctg caccocgct tcgacttctt cgcgggcgtg 1680
gccgacgtcg agcttcccg cggcgaaagtc ccccgccgcg gtccgggggc gatccaggcc 1740
acctggcgcg tggtagcagg caacctgccc ctggcgctgt gtccgggtggc gtttctgtac 1800
gcccgggggc tggagctcgg cgttggccgc cagcccatgg cgccggctac catagccgcc 1860
gtccgcgggg cgttcgagga ccgcagctac ccggcggtgt tttacctgct gcaagccgcg 1920
attcacggca gcgagcacgt gttctgcgcc ctggcgcggc tcgtgactca gtgcatcacc 1980
agctactgga acaacacgcg atgcgcggcg ttctgtaacg actactcgtt ggtctcgtac 2040
atcgtgacct acctcggggg cgacctcccc gaggagtga tggcctgta tcgggacctg 2100
gtggccacag tcgaggccct ggcccagctg gtggacgact ttacctgce gggcccgag 2160
ctgggcgggg aggtcaggc cgagctgaat cacctgatgc gcgaccggc gctgctgcg 2220
cccctcgtgt gggactgca cgcccttatg cgacacgcgg ccctggaccg ccaccgagac 2280
tgccgattg acgcgggggg gcacgagccc gtctacgcgg cggcgtgcaa cgtggcgacg 2340
gccgacttta accgcaacga cggccggctg ctgcacaaca cccaggcccg cgcggtcgac 2400
gccgcccagc accggccgca ccggccggcc gactggaccg tccaccaca aatctactat 2460
tacgtgctgg tgccggcctt ctgcggggg cgtgctgca ccgcggggt ccgcttcgac 2520
cgcgtgtacg ccacgtgca gaacatggtg gtcccgaga tcgccccgg tgaggagtgc 2580
ccgagcgatc ccgtgacga ccccgccac ccgctgcac cgcgaatct ggtggccaac 2640
acggtcaacg ccatgttcca caacgggcgc gtcgtcgtcg acgggcccgc catgctcacg 2700
ctgcaggtgc tggcgcaaaa catggccgag cgcacgacgg cgtgctgtg ctcccgcgcg 2760
cccgcgcgg gcgccaacac cgcgtcgac gccaacatgc gcatcttga cggggcgctg 2820
cacgcccgg tgcgtgcat ggccccccag cacctggacc acaccatcca aaatggcgaa 2880
tacttctacg tcttggccgt ccacgcgctg ttgctggcg ccgaccagt ggccaacgcg 2940
cccaacttcc ccccgccct gcgcgacctg gcgcgccag tccccctggt ccccccggcc 3000
ctgggggcca actacttctc ctccatccgc cagcccgtgg tgcagcacgc ccgcgagagc 3060
gcggcggggg agaacgcgct gacctacgcg ctcatggcg ggtacttcaa gatgagcccc 3120
gtggccctgt atcaccagct caagacgggc ctccaccccg ggttcgggtt caccgtcgtg 3180

```

```

cggcaggacc gcttcgtgac cgagaacgtg ctgttttccg agcgcgcgtc ggaggcgtac 3240
tttctgggcc agctccaggt ggcccgccac gaaacgggcg ggggggtcag cttcacgtc 3300
accagccgc gcggaacgt ggacctgggt gtgggctaca ccgccgtcg gccacggcc 3360
accgtccgca accccgttac ggacatgggc aacctcccc aaaactttta cctcgccgc 3420
ggggccccc cgctgctaga caacgcggcc gccgtgtacc tgcgcaacgc ggtcgtggcg 3480
ggaacccggc tggggccggc ccagcccctc ccggtctttg gctgcgccca ggtgcccgcg 3540
cgcccgggca tggaccacgg gcaggatgcc gtgtgtgagt tcatcgccac ccccggtggc 3600
acggacatca actactttcg ccggccctgc aaccgcggg gacgcgcggc cggcgcggtg 3660
tacgcggggg acaaggaggg ggacgtcata gccctcatgt acgaccacgg ccagagcgac 3720
ccggcgcggc ccttcgcggc cacggccaac ccgtgggctg cgcagcgggt ctctacggc 3780
gacctgctgt acaacggggc ctatcacctc aacggggcct cggccgtcct cagcccctgc 3840
ttcaagtct taccgcggc cgacatcacg gccaaacatc gctgcctgga gcgtcttct 3900
gtggaaacgg gatcgcggt atccacggcc accgtgccg gcgacgtgca gtttaagcg 3960
ccggcggggt gccgcgagct cgtggaagac ccgtgcggcc tgtttcagga agcctaccg 4020
atcacctgcg ccagcgaccc cgccctgcta cgcagcgccc gcgatgggga ggccacgcg 4080
cgagagaccc actttacgca gtatctcatc tacgacgct ccccgctaaa gggcctgtct 4140
ctgtaatga 4149

```

<210> 85

<211> 1623

<212> DNA

<213> Homo sapiens

<400> 85

```

atgcagcatc accaccatca ccacgagctc agctatgcca ccaccctgca ccaccgggac 60
gttgtgtttt acgtcacggc agacagaaac cgcgcctact ttgtgtgcgg ggggtccgtt 120
tattccgtag ggcggcctcg ggattctcag ccgggggaaa ttgccaagt tggcctggtg 180
gtccggggga caggcccaa agaccgcatg gtcgccaact acgtacgaag cgagctccgc 240
cagcgcgggc tgcgggacgt gcggcccgtg ggggaggacg aggtgttcct ggacagcgtg 300
tgtctgctaa acccgaacgt gagctccgag cgagacgtga ttaataccaa cgacgttgaa 360
gtgctggacg aatgcctggc cgaatactgc acctcgctgc gaaccagccc ggggtgctg 420
gtgacggggg tgcgcgtgcg cgcgcgagac aggttcactc agctatttga gcaccggcg 480
atcgtaaca tttcctcgct cttcgcgtac acccctccc cctacgtatt cgccctggcc 540
caggcgaccc tccccggct cccgagctcg ctggagcccc tggtagcgcg cctgtttgac 600
ggcattcccg ccccgcgcca gccctggac gcccgcgacc ggcgcacgga tgtcgtgatc 660
acgggcaccc gcgccccag accgatggcc gggaccgggg ccgggggcgc gggggccaag 720
cgggccaccg tcagcgagtt cgtgcaagtg aagcacatcg accgtgttgt gtccccgagc 780
gtctcttccg ccccccgcc gagcgcccc gacgcgagtc tgccgcccc ggggtccag 840
gaggcgcccc cgccgggccc cccgctcagg gagctgtggt ggggtgttcta cgccggcgac 900
cgggcgctgg aggagcccca cgccgagtcg ggattgacgc gcgaggaggt ccgcgccgtg 960
catgggttcc gggagcaggc gtggaagctg tttgggtcgg tgggggtccc gcggcggtt 1020
ctcggggcgg cgctggccct gagcccgacc caaaagctcg ccgtctacta ctatctcatc 1080
caccgggagc ggcgcagtgc ccccttcccc gcgtcgtgc ggctcgtcg tcggtacatc 1140
cagcgccacg gcctgtacgt tcccgcgccc gacgaaccga cgttggccga tgccatgaac 1200
gggctgttcc gcgacgcgct ggcgccggg accgtggccg agcagctcct catgttcgac 1260
ctcctccgc ccaaggacgt gccgtgggg agcgacgcgc gggccgacag cgccgccctg 1320
ctgcgctttg tggactcgca acgctgacc ccgggggggt ccgtctcgcc cgagcacgtc 1380
atgtacctcg gcgcgttcct gggcgtgttg tacgcggccc acggacgcct ggcccgggc 1440
acgcataccg cgcgcctgac gggcgtgacg tccctggtcc tgaccgtggg ggacgtcgac 1500
cggatgtccg cgtttgaccg cgggcccggc ggggcggtg gccgcacgcg aaccgcggg 1560
tacctggacg cgctgcttac cgtttgcctg gctcgcgccc agcacggcca gtctgtgtga 1620
tga 1623

```

<210> 86

<211> 2211

<212> DNA

<213> Homo sapiens

<400> 86

```

atgcagcatc accaccatca ccacgtgtct atcgaaggtc gtgctagctc tgggtggcagc 60
ggtctggttc cgcgtggttag ctctggttcg ggggacgacg acgacaaatc tagtaggcac 120
tccgtgcgcg ggcattgccgt acgccggagg cgcgcctcca cccggtccca tgccccgtcc 180
gcgcattcgc cgcactcgcc cgtggaggac gagccccagg gcggtggagt cgggttaatg 240
gggtacctgc gggcgggtgtt taacgtggac gacgacagcg aggtcgaggc cgcgggggag 300
atggcgagcg aagagccgcc cccgcgcgt cgcgggagg cccgcggtca ccccggtcc 360
cgacgcgcgt ccgagggccc ggccggcgcg ccccccgcg gggcgctcct tccgcgcccc 420
aggtccgtta cggccaggag ccagtcggtt cgcggacgcc gggacagcgc catcacgcgg 480
gccccgcggg gaggctacct gggcccgatg gacccacgcg acgttttggg gcgggtgggc 540
ggttcgcggg tgggtgccctc gccgctgttc ctggacgagc tcagctacga ggaggacgac 600
taccgcgcgc cgtcgcgcga cgatgacggc gccggggcgc ggccctccgc gacggtcgag 660
attctcgcgg gccgcgtgtc gggcccgagg ctgcaggcgc cattccccct ggaccgcctg 720
accccccgag tcgcgcgtg ggacgagtc gtgcgctcgg ccctggccct gggacatccg 780
gcggggttct acccgtgtcc ggatagcgcg ttccgggtgt cgcgcgtggg ggtcatgcac 840
tttgccctcc cggccgaccc aaagggtgtt ttccgccaga cgtgcagca gggcgaggcg 900
ctggcctggt acgtcacggg cgacgcgac ctgcacctga cggatcggcg ggcaaaaacc 960
agccccctcc gcgcgatggg ttttctggtg gacgccatcg tgcgggtggc gatcaacggg 1020
tgggtctgcg ggacgcgcct gcacacggag ggcgcggct cggagctcga cgacaggcg 1080
gccgagctcc gacggcagtt cgcgagctc acggcgttgc ggcccggtgg ggcgcgcgc 1140
gtgccgctgc tcagcgcggg aggggcgcgc cccccccacc ccggccccga cgcgcggtc 1200
tttcgcagtt cgtggggtc cctgctgtac tggcccgggg tgcgcgcgt cctggggcgc 1260
gactgtcgcg tggccgcccg ctacgcgggg cgcattgacg acatcgccac cggggctctg 1320
ctgcgccgct tcaaccccgc gcccgtaaaa tgcgtgctcc cgcgggaggc cgcgtttgcg 1380
gggcgcgtcc tggacgtgct ggccggtcct gcggagcaga cgggtccagt gctctcgtg 1440
gtcgtggggg cgcgcctgca cccgcactcc gccaccccgc cgtttgcgga cgtggagcag 1500
gaggcgtgt ttgcgcctt gccctgggc agccccggg tcgtggcggc cgagcacgag 1560
gcgctgggcg acaccgcgc gcgcgcctc ctgcacca gcgggtgaa cgcgtgctg 1620
ggcgcgccgc tgtacgcgt gcacacggc ctggcgacc ttacctgaa ataccctg 1680
gcctgcgggg acgcgcgcgc gcgcaggac gacgcgcgc ccgcgcgcgc cgtgctggc 1740
acggggctca tctgcagcg gctgctggg ctggccgaca cgggtggcgc gtgcgtggc 1800
ctggccgcgt ttgacggcg gtcgacggc cccgaggtg gcacgtacac cccctgcgc 1860
tacgcgtgcg tctccgcgc gacccagccc ctgtacgcgc ggaccacccc cgccaaattt 1920
tgggcggacg tgcgcgcgc gcgggaacac gtggacctc gccccgcgc ctcggcgcgc 1980
cgggcgcgcg tgagcgggac ggcagacccc gccttctgc tcgaagacct ggcggcctc 2040
ccccgcgc ccctgaatag cgagtcgtg ctggggccgc ggggtccggt cgtggacatc 2100
atggcgagcgt ttccgaaact gctcatggg gacgaggaga ccgcgcctt ccgggcgcac 2160
gtgtccggga ggcgcgcgac cgggctgggc ggcccgcac gccatagtg a 2211

```

<210> 87

<211> 2118

<212> DNA

<213> Homo sapiens

<400> 87

```

atgcagcatc accaccatca ccaccactcc gtgcgcgggc atgccgtacg ccggaggcgc 60
gcctccaccc ggtcccatgc cccgtccgcg catcgcgcgc actcgcgcgt ggaggacgag 120
cccgagggcg gtggaagtcg gttaatggg tacctgcggg cgggtgttaa cgtggacgac 180
gacagcgagg tcgaggccgc gggggagatg gcgagcgaag agccgcccc gcgccgtgc 240
cgggaggccc gcggtcacc cgggtcccga cgcgcgtccg agggccgggc ggccggcgc 300
ccccgcggg cgtcctttcc gcgcccagg tccgttaagg ccaggagcca gtccgttcgc 360
ggacgcggg acagcgcct cagcggggc ccgcggggg gctacctggg ccgatggac 420
ccacgcgacg tttggggcg ggtgggcgt tcgcgggtg tgccctcgc gctgttcctg 480
gacgagctca gctacgagga ggaacgtac cccgcgcgc tcgcgcacga tgacggcgc 540
ggggcgcggc ctccgcgcg ggtcgagatt ctgcggggc gcgtgtcggg ccggagctg 600
caggcggcat tccccctgga ccgctgacc ccccgagtc ccgcgtggga cgagtcctg 660
cgctcggccc tggccctggg acatccggc ggggttctac cgtgtccgga tagcgcgtt 720
gggctgtcgc gcgtgggggt catgcacttt gcctccccgc ccgacccaaa ggtgtttttc 780

```

```

cgccagacgc tgcagcaggg cgagggcgctg gectggtagc tcacgggcga cgcgatactc 840
gacctgacgg atcggcgggc aaaaaccagc cctcccgcg cgatgggttt tctgggtggac 900
gccatcgtgc ggggtggcgat caacgggtgg gtctgcggga cgcgcctgca cacggagggg 960
cgcggctcgg agctcgacga cagggcggcc gagctccgac ggcagttcgc gaggctcacg 1020
gcgttgccgg ccgtgggggc cgcgcctg cgcgtgctca gcgcgggagg ggccgcgccc 1080
ccccaccccg gccccgacgc cgcggtcttt cgcagttcgc tggggtccct gctgtactgg 1140
cccggggtgc gcgcgctcct ggggcgcgac tgcgcgctgg ccgcccgcta cgcggggcgc 1200
atgacgtaca tcgccaccgg ggctctgctc gcccgttca accccggcgc cgtcaaattgc 1260
gtgctcccgc gggaggccgc gtttgccggg cgcgtcctgg acgtgctggc ggtcctggcg 1320
gagcagacgg tccagtggct ctcggtggct gtggggggcg gcctgcaccc gcaactccgc 1380
caccgccgct ttgcggacgt ggagcaggag gcgctgtttc gcgccctgcc cctgggcagc 1440
cccggggtcg tggcgccga gcacgaggcg ctgggcgaca ccgcggcgcg ccgcctgctc 1500
gccaccagcg ggtgaacgc cgtgctgggc gcggccgtgt acgcgctgca cacggccctg 1560
gcgaccgtta ccctgaaata cgcctggcc tgcggggacg cgcgccggcg cagggaacgac 1620
gcggcgccgc cgcgcgccgt gctggcgacg gggctcatcc tgcagcggt gctgggcctg 1680
gccgacacgg tggtcgcgtg cgtggccctg gccgcgtttg acggcgggtc gacggccccc 1740
gaggtgggca cgtacacccc cctgcgtac gcgtgcgtcc tccgcgcgac ccagcccctg 1800
tacgcgcgga ccacccccc caaatttttg gcggacgtgc gcgccgcgc ggaacacgtg 1860
gaccttcgcc ccgcgtcctc ggcccccg gcgcccgta gcgggacggc agaccccgcc 1920
ttcctgctcg aagacctggc ggccttcccc cccgcccccc tgaatagcga gtccgtgctg 1980
gggcgcgggg tccgcgtcgt ggacatcatg gcgcagtttc ggaaactgct catgggcgac 2040
gaggagaccg ccgccctccg ggcgcacgtg tccgggaggc gcgcgaccgg gctgggcggc 2100
ccgccacgcc catagtga

```

<210> 88

<211> 939

<212> DNA

<213> Homo sapiens

<400> 88

```

atgcagcatc accaccatca ccaccaactcc gtgcgcgggc atgccgtacg ccggaggcgc 60
gctccacccc ggtcccacgc cccgtccgog catcgcgcgc actcggccgt ggaggacgag 120
cccgagggcg gtggagtcgg gttaatgggg tacctgcggg cgggtgttaa cgtggacgac 180
gacagtgagg tcgagggcgc gggggagatg gcgagcgaag agccgcccc gcgcgctcgc 240
cgggaggccc gcggtcaccc cgggtccoga cgcgcgtccg aggccgggc ggccggccc 300
ccccgcgggg cgtcctttcc gcgcccagg tccgttacgg ccaggagcca gtccgttcgc 360
ggacgcgggg acagcgccat cacgcgggcc ccgcggggag gctacctggg cccgatggac 420
ccacgcgacg ttttgggcg ggtgggcggg tcgcgggtgg tgcctcgc gctgttctc 480
gacgagctca gctacgagga ggacgactac cccgcgcgc tcgcgcacga tgacggcgcc 540
ggggcgcggc ctcccgcgac ggtcgagatt ctgcggggc gcgtgtcggg cccggagctg 600
caggcgcat tccccctgga ccgcctgacc ccccgagtcg ccgcgtggga caggtccgtg 660
cgctcggccc tggccctggg acatccggcc gggttctacc cgtgtccgga tagcgcgtt 720
gggctgtcgc gcgtgggggt catgcacttt gcctccccg ccgacccaaa ggtgttttt 780
cgccagacgc tgcagcagg cgagggcgtg gctggtagc tcacgggcga cgcgatactc 840
gacctgacgg atcggcgggc aaaaaccagc cctcccgcg cgatgggctt tctgggtggac 900
gccatcgtgc ggggtggcgat caacgggtgg gtctgatga

```

<210> 89

<211> 843

<212> DNA

<213> Homo sapiens

<400> 89

```

atgcagcatc accaccatca ccaccacgcc gccgcacccc aggcggacgt ggcccggtt 60
ctggacagcc agccactgt gggaacggac cccggctacc cagtcacctt agaactcacg 120
cccgagaacg cggaggcggg ggcgcgggtt ctgggggacg ccgtcgaccg cgagcccgcg 180
ctcatgctgg agtacttctg tcggtgcgcc cgcgaggaga gcaagcgcgt gccccacga 240
accttcggca gcgccccccg cctcacggag gacgactttg ggctcctgaa ctacgcgctc 300

```

```

gctgagatgc gacgcctgtg cctggacott cccccggtcc cccccaacgc atacacgccc 360
tatcatctga gggagtatgc gacgcggctg gttaacgggt tcaaaccctt ggtgcggcgg 420
tccgcccgcc tgtatcgcat cctggggatt ctggttcacc tgcgcatccg taccggggag 480
gcctcctttg aggaatggat gcgctccaag gaggtggacc tggacttcgg gctgacggaa 540
aggcttcgcg aacacgaggc ccagctaatt atcctggccc aggccctgaa cccctacgac 600
tgtctgatcc acagcacccc gaacacgctc gtcgagcggg ggctgcagtc ggcgctgaag 660
tacgaagagt ttacacctaa gcgcttcggc gggcactaca tggagtccgt cttccagatg 720
tacacccgca tcgccgggtt cctggcgctg cgggcgaccc gcggcatgcg ccacatcgcc 780
ctggggcgac aggggtcgtg gtgggaaatg ttcaagttct tttccaccg cctctactaa 840
tga 843

```

<210> 90

<211> 279

<212> PRT

<213> Homo sapiens

<400> 90

```

Met Gln His His His His His His His Ala Ala Ala Pro Gln Ala Asp
                    5                      10                      15

```

```

Val Ala Pro Val Leu Asp Ser Gln Pro Thr Val Gly Thr Asp Pro Gly
                20                      25                      30

```

```

Tyr Pro Val Pro Leu Glu Leu Thr Pro Glu Asn Ala Glu Ala Val Ala
                35                      40                      45

```

```

Arg Phe Leu Gly Asp Ala Val Asp Arg Glu Pro Ala Leu Met Leu Glu
                50                      55                      60

```

```

Tyr Phe Cys Arg Cys Ala Arg Glu Glu Ser Lys Arg Val Pro Pro Arg
                65                      70                      75                      80

```

```

Thr Phe Gly Ser Ala Pro Arg Leu Thr Glu Asp Asp Phe Gly Leu Leu
                85                      90                      95

```

```

Asn Tyr Ala Leu Ala Glu Met Arg Arg Leu Cys Leu Asp Leu Pro Pro
                100                      105                      110

```

```

Val Pro Pro Asn Ala Tyr Thr Pro Tyr His Leu Arg Glu Tyr Ala Thr
                115                      120                      125

```

```

Arg Leu Val Asn Gly Phe Lys Pro Leu Val Arg Arg Ser Ala Arg Leu
                130                      135                      140

```

```

Tyr Arg Ile Leu Gly Ile Leu Val His Leu Arg Ile Arg Thr Arg Glu
                145                      150                      155                      160

```

```

Ala Ser Phe Glu Glu Trp Met Arg Ser Lys Glu Val Asp Leu Asp Phe
                165                      170                      175

```

```

Gly Leu Thr Glu Arg Leu Arg Glu His Glu Ala Gln Leu Met Ile Leu
                180                      185                      190

```

```

Ala Gln Ala Leu Asn Pro Tyr Asp Cys Leu Ile His Ser Thr Pro Asn
                195                      200                      205

```

```

Thr Leu Val Glu Arg Gly Leu Gln Ser Ala Leu Lys Tyr Glu Glu Phe
                210                      215                      220

```

Tyr Leu Lys Arg Phe Gly Gly His Tyr Met Glu Ser Val Phe Gln Met
 225 230 235 240

Tyr Thr Arg Ile Ala Gly Phe Leu Ala Cys Arg Ala Thr Arg Gly Met
 245 250 255

Arg His Ile Ala Leu Gly Arg Gln Gly Ser Trp Trp Glu Met Phe Lys
 260 265 270

Phe Phe Phe His Arg Leu Tyr
 275

<210> 91

<211> 539

<212> PRT

<213> Homo sapiens

<400> 91

Met Gln His His His His His His Glu Leu Ser Tyr Ala Thr Thr Leu
 5 10 15

His His Arg Asp Val Val Phe Tyr Val Thr Ala Asp Arg Asn Arg Ala
 20 25 30

Tyr Phe Val Cys Gly Gly Ser Val Tyr Ser Val Gly Arg Pro Arg Asp
 35 40 45

Ser Gln Pro Gly Glu Ile Ala Lys Phe Gly Leu Val Val Arg Gly Thr
 50 55 60

Gly Pro Lys Asp Arg Met Val Ala Asn Tyr Val Arg Ser Glu Leu Arg
 65 70 75 80

Gln Arg Gly Leu Arg Asp Val Arg Pro Val Gly Glu Asp Glu Val Phe
 85 90 95

Leu Asp Ser Val Cys Leu Leu Asn Pro Asn Val Ser Ser Glu Arg Asp
 100 105 110

Val Ile Asn Thr Asn Asp Val Glu Val Leu Asp Glu Cys Leu Ala Glu
 115 120 125

Tyr Cys Thr Ser Leu Arg Thr Ser Pro Gly Val Leu Val Thr Gly Val
 130 135 140

Arg Val Arg Ala Arg Asp Arg Val Ile Glu Leu Phe Glu His Pro Ala
 145 150 155 160

Ile Val Asn Ile Ser Ser Arg Phe Ala Tyr Thr Pro Ser Pro Tyr Val
 165 170 175

Phe Ala Leu Ala Gln Ala His Leu Pro Arg Leu Pro Ser Ser Leu Glu
 180 185 190

Pro Leu Val Ser Gly Leu Phe Asp Gly Ile Pro Ala Pro Arg Gln Pro
 195 200 205

Leu Asp Ala Arg Asp Arg Arg Thr Asp Val Val Ile Thr Gly Thr Arg
 210 215 220
 Ala Pro Arg Pro Met Ala Gly Thr Gly Ala Gly Gly Ala Gly Ala Lys
 225 230 235 240
 Arg Ala Thr Val Ser Glu Phe Val Gln Val Lys His Ile Asp Arg Val
 245 250 255
 Val Ser Pro Ser Val Ser Ser Ala Pro Pro Pro Ser Ala Pro Asp Ala
 260 265 270
 Ser Leu Pro Pro Pro Gly Leu Gln Glu Ala Ala Pro Pro Gly Pro Pro
 275 280 285
 Leu Arg Glu Leu Trp Trp Val Phe Tyr Ala Gly Asp Arg Ala Leu Glu
 290 295 300
 Glu Pro His Ala Glu Ser Gly Leu Thr Arg Glu Glu Val Arg Ala Val
 305 310 315 320
 His Gly Phe Arg Glu Gln Ala Trp Lys Leu Phe Gly Ser Val Gly Ala
 325 330 335
 Pro Arg Ala Phe Leu Gly Ala Ala Leu Ala Leu Ser Pro Thr Gln Lys
 340 345 350
 Leu Ala Val Tyr Tyr Tyr Leu Ile His Arg Glu Arg Arg Met Ser Pro
 355 360 365
 Phe Pro Ala Leu Val Arg Leu Val Gly Arg Tyr Ile Gln Arg His Gly
 370 375 380
 Leu Tyr Val Pro Ala Pro Asp Glu Pro Thr Leu Ala Asp Ala Met Asn
 385 390 395 400
 Gly Leu Phe Arg Asp Ala Leu Ala Ala Gly Thr Val Ala Glu Gln Leu
 405 410 415
 Leu Met Phe Asp Leu Leu Pro Pro Lys Asp Val Pro Val Gly Ser Asp
 420 425 430
 Ala Arg Ala Asp Ser Ala Ala Leu Leu Arg Phe Val Asp Ser Gln Arg
 435 440 445
 Leu Thr Pro Gly Gly Ser Val Ser Pro Glu His Val Met Tyr Leu Gly
 450 455 460
 Ala Phe Leu Gly Val Leu Tyr Ala Gly His Gly Arg Leu Ala Ala Ala
 465 470 475 480
 Thr His Thr Ala Arg Leu Thr Gly Val Thr Ser Leu Val Leu Thr Val
 485 490 495
 Gly Asp Val Asp Arg Met Ser Ala Phe Asp Arg Gly Pro Ala Gly Ala
 500 505 510

Ala Gly Arg Thr Arg Thr Ala Gly Tyr Leu Asp Ala Leu Leu Thr Val
 515 520 525

Cys Leu Ala Arg Ala Gln His Gly Gln Ser Val
 530 535

<210> 92
 <211> 858
 <212> PRT
 <213> Homo sapiens

<400> 92
 Met Gln His His His His His His Met Ser Asp Lys Ile Ile His Leu
 5 10 15

Thr Asp Asp Ser Phe Asp Thr Asp Val Leu Lys Ala Asp Gly Ala Ile
 20 25 30

Leu Val Asp Phe Trp Ala Glu Trp Cys Gly Pro Cys Lys Met Ile Ala
 35 40 45

Pro Ile Leu Asp Glu Ile Ala Asp Glu Tyr Gln Gly Lys Leu Thr Val
 50 55 60

Ala Lys Leu Asn Ile Asp Gln Asn Pro Gly Thr Ala Pro Lys Tyr Gly
 65 70 75 80

Ile Arg Gly Ile Pro Thr Leu Leu Leu Phe Lys Asn Gly Glu Val Ala
 85 90 95

Ala Thr Lys Val Gly Ala Leu Ser Lys Gly Gln Leu Lys Glu Phe Leu
 100 105 110

Asp Ala Asn Leu Ala Gly Ser Gly Ser Gly His Met Gln His His His
 115 120 125

His His His Val Ser Ile Glu Gly Arg Ala Ser Ser Gly Gly Ser Gly
 130 135 140

Leu Val Pro Arg Gly Ser Ser Gly Ser Gly Asp Asp Asp Asp Lys Ser
 145 150 155 160

Ser Arg His Ser Val Arg Gly His Ala Val Arg Arg Arg Arg Ala Ser
 165 170 175

Thr Arg Ser His Ala Pro Ser Ala His Arg Ala Asp Ser Pro Val Glu
 180 185 190

Asp Glu Pro Glu Gly Gly Gly Val Gly Leu Met Gly Tyr Leu Arg Ala
 195 200 205

Val Phe Asn Val Asp Asp Asp Ser Glu Val Glu Ala Ala Gly Glu Met
 210 215 220

Ala Ser Glu Glu Pro Pro Pro Arg Arg Arg Arg Glu Ala Arg Gly His
 225 230 235 240

Pro Gly Ser Arg Arg Ala Ser Glu Ala Arg Ala Ala Ala Pro Pro Arg
 245 250 255
 Arg Ala Ser Phe Pro Arg Pro Arg Ser Val Thr Ala Arg Ser Gln Ser
 260 265 270
 Val Arg Gly Arg Arg Asp Ser Ala Ile Thr Arg Ala Pro Arg Gly Gly
 275 280 285
 Tyr Leu Gly Pro Met Asp Pro Arg Asp Val Leu Gly Arg Val Gly Gly
 290 295 300
 Ser Arg Val Val Pro Ser Pro Leu Phe Leu Asp Glu Leu Ser Tyr Glu
 305 310 315 320
 Glu Asp Asp Tyr Pro Ala Ala Val Ala His Asp Asp Gly Ala Gly Ala
 325 330 335
 Arg Pro Pro Ala Thr Val Glu Ile Leu Ala Gly Arg Val Ser Gly Pro
 340 345 350
 Glu Leu Gln Ala Ala Phe Pro Leu Asp Arg Leu Thr Pro Arg Val Ala
 355 360 365
 Ala Trp Asp Glu Ser Val Arg Ser Ala Leu Ala Leu Gly His Pro Ala
 370 375 380
 Gly Phe Tyr Pro Cys Pro Asp Ser Ala Phe Gly Leu Ser Arg Val Gly
 385 390 395 400
 Val Met His Phe Ala Ser Pro Ala Asp Pro Lys Val Phe Phe Arg Gln
 405 410 415
 Thr Leu Gln Gln Gly Glu Ala Leu Ala Trp Tyr Val Thr Gly Asp Ala
 420 425 430
 Ile Leu Asp Leu Thr Asp Arg Arg Ala Lys Thr Ser Pro Ser Arg Ala
 435 440 445
 Met Gly Phe Leu Val Asp Ala Ile Val Arg Val Ala Ile Asn Gly Trp
 450 455 460
 Val Cys Gly Thr Arg Leu His Thr Glu Gly Arg Gly Ser Glu Leu Asp
 465 470 475 480
 Asp Arg Ala Ala Glu Leu Arg Arg Gln Phe Ala Ser Leu Thr Ala Leu
 485 490 495
 Arg Pro Val Gly Ala Ala Ala Val Pro Leu Leu Ser Ala Gly Gly Ala
 500 505 510
 Ala Pro Pro His Pro Gly Pro Asp Ala Ala Val Phe Arg Ser Ser Leu
 515 520 525
 Gly Ser Leu Leu Tyr Trp Pro Gly Val Arg Ala Leu Leu Gly Arg Asp
 530 535 540
 Cys Arg Val Ala Ala Arg Tyr Ala Gly Arg Met Thr Tyr Ile Ala Thr

545	550	555	560
Gly Ala Leu Leu Ala Arg Phe Asn Pro Gly Ala Val Lys Cys Val Leu	565	570	575
Pro Arg Glu Ala Ala Phe Ala Gly Arg Val Leu Asp Val Leu Ala Val	580	585	590
Leu Ala Glu Gln Thr Val Gln Trp Leu Ser Val Val Val Gly Ala Arg	595	600	605
Leu His Pro His Ser Ala His Pro Ala Phe Ala Asp Val Glu Gln Glu	610	615	620
Ala Leu Phe Arg Ala Leu Pro Leu Gly Ser Pro Gly Val Val Ala Ala	625	630	635
Glu His Glu Ala Leu Gly Asp Thr Ala Ala Arg Arg Leu Leu Ala Thr	645	650	655
Ser Gly Leu Asn Ala Val Leu Gly Ala Ala Val Tyr Ala Leu His Thr	660	665	670
Ala Leu Ala Thr Val Thr Leu Lys Tyr Ala Leu Ala Cys Gly Asp Ala	675	680	685
Arg Arg Arg Arg Asp Asp Ala Ala Ala Ala Arg Ala Val Leu Ala Thr	690	695	700
Gly Leu Ile Leu Gln Arg Leu Leu Gly Leu Ala Asp Thr Val Val Ala	705	710	715
Cys Val Ala Leu Ala Ala Phe Asp Gly Gly Ser Thr Ala Pro Glu Val	725	730	735
Gly Thr Tyr Thr Pro Leu Arg Tyr Ala Cys Val Leu Arg Ala Thr Gln	740	745	750
Pro Leu Tyr Ala Arg Thr Thr Pro Ala Lys Phe Trp Ala Asp Val Arg	755	760	765
Ala Ala Ala Glu His Val Asp Leu Arg Pro Ala Ser Ser Ala Pro Arg	770	775	780
Ala Pro Val Ser Gly Thr Ala Asp Pro Ala Phe Leu Leu Glu Asp Leu	785	790	795
Ala Ala Phe Pro Pro Ala Pro Leu Asn Ser Glu Ser Val Leu Gly Pro	805	810	815
Arg Val Arg Val Val Asp Ile Met Ala Gln Phe Arg Lys Leu Leu Met	820	825	830
Gly Asp Glu Glu Thr Ala Ala Leu Arg Ala His Val Ser Gly Arg Arg	835	840	845
Ala Thr Gly Leu Gly Gly Pro Pro Arg Pro	850	855	

<210> 93
 <211> 311
 <212> PRT
 <213> Homo sapiens

<400> 93

```

Met Gln His His His His His His His Ser Val Arg Gly His Ala Val
              5                      10                      15

Arg Arg Arg Arg Ala Ser Thr Arg Ser His Ala Pro Ser Ala His Arg
              20                      25                      30

Ala Asp Ser Pro Val Glu Asp Glu Pro Glu Gly Gly Gly Val Gly Leu
              35                      40                      45

Met Gly Tyr Leu Arg Ala Val Phe Asn Val Asp Asp Asp Ser Glu Val
              50                      55                      60

Glu Ala Ala Gly Glu Met Ala Ser Glu Glu Pro Pro Pro Arg Arg Arg
              65                      70                      75                      80

Arg Glu Ala Arg Gly His Pro Gly Ser Arg Arg Ala Ser Glu Ala Arg
              85                      90                      95

Ala Ala Ala Pro Pro Arg Arg Ala Ser Phe Pro Arg Pro Arg Ser Val
              100                      105                      110

Thr Ala Arg Ser Gln Ser Val Arg Gly Arg Arg Asp Ser Ala Ile Thr
              115                      120                      125

Arg Ala Pro Arg Gly Gly Tyr Leu Gly Pro Met Asp Pro Arg Asp Val
              130                      135                      140

Leu Gly Arg Val Gly Gly Ser Arg Val Val Pro Ser Pro Leu Phe Leu
              145                      150                      155                      160

Asp Glu Leu Ser Tyr Glu Glu Asp Asp Tyr Pro Ala Ala Val Ala His
              165                      170                      175

Asp Asp Gly Ala Gly Ala Arg Pro Pro Ala Thr Val Glu Ile Leu Ala
              180                      185                      190

Gly Arg Val Ser Gly Pro Glu Leu Gln Ala Ala Phe Pro Leu Asp Arg
              195                      200                      205

Leu Thr Pro Arg Val Ala Ala Trp Asp Glu Ser Val Arg Ser Ala Leu
              210                      215                      220

Ala Leu Gly His Pro Ala Gly Phe Tyr Pro Cys Pro Asp Ser Ala Phe
              225                      230                      235                      240

Gly Leu Ser Arg Val Gly Val Met His Phe Ala Ser Pro Ala Asp Pro
              245                      250                      255

Lys Val Phe Phe Arg Gln Thr Leu Gln Gln Gly Glu Ala Leu Ala Trp
              260                      265                      270

```

Tyr Val Thr Gly Asp Ala Ile Leu Asp Leu Thr Asp Arg Arg Ala Lys
 275 280 285

Thr Ser Pro Ser Arg Ala Met Gly Phe Leu Val Asp Ala Ile Val Arg
 290 295 300

Val Ala Ile Asn Gly Trp Val
 305 310

<210> 94

<211> 704

<212> PRT

<213> Homo sapiens

<400> 94

Met Gln His His His His His His His Ser Val Arg Gly His Ala Val
 5 10 15

Arg Arg Arg Arg Ala Ser Thr Arg Ser His Ala Pro Ser Ala His Arg
 20 25 30

Ala Asp Ser Pro Val Glu Asp Glu Pro Glu Gly Gly Gly Val Gly Leu
 35 40 45

Met Gly Tyr Leu Arg Ala Val Phe Asn Val Asp Asp Asp Ser Glu Val
 50 55 60

Glu Ala Ala Gly Glu Met Ala Ser Glu Glu Pro Pro Pro Arg Arg Arg
 65 70 75 80

Arg Glu Ala Arg Gly His Pro Gly Ser Arg Arg Ala Ser Glu Ala Arg
 85 90 95

Ala Ala Ala Pro Pro Arg Arg Ala Ser Phe Pro Arg Pro Arg Ser Val
 100 105 110

Thr Ala Arg Ser Gln Ser Val Arg Gly Arg Arg Asp Ser Ala Ile Thr
 115 120 125

Arg Ala Pro Arg Gly Gly Tyr Leu Gly Pro Met Asp Pro Arg Asp Val
 130 135 140

Leu Gly Arg Val Gly Gly Ser Arg Val Val Pro Ser Pro Leu Phe Leu
 145 150 155 160

Asp Glu Leu Ser Tyr Glu Glu Asp Asp Tyr Pro Ala Ala Val Ala His
 165 170 175

Asp Asp Gly Ala Gly Ala Arg Pro Pro Ala Thr Val Glu Ile Leu Ala
 180 185 190

Gly Arg Val Ser Gly Pro Glu Leu Gln Ala Ala Phe Pro Leu Asp Arg
 195 200 205

Leu Thr Pro Arg Val Ala Ala Trp Asp Glu Ser Val Arg Ser Ala Leu
 210 215 220

Ala Leu Gly His Pro Ala Gly Phe Tyr Pro Cys Pro Asp Ser Ala Phe
 225 230 235 240
 Gly Leu Ser Arg Val Gly Val Met His Phe Ala Ser Pro Ala Asp Pro
 245 250 255
 Lys Val Phe Phe Arg Gln Thr Leu Gln Gln Gly Glu Ala Leu Ala Trp
 260 265 270
 Tyr Val Thr Gly Asp Ala Ile Leu Asp Leu Thr Asp Arg Arg Ala Lys
 275 280 285
 Thr Ser Pro Ser Arg Ala Met Gly Phe Leu Val Asp Ala Ile Val Arg
 290 295 300
 Val Ala Ile Asn Gly Trp Val Cys Gly Thr Arg Leu His Thr Glu Gly
 305 310 315 320
 Arg Gly Ser Glu Leu Asp Asp Arg Ala Ala Glu Leu Arg Arg Gln Phe
 325 330 335
 Ala Ser Leu Thr Ala Leu Arg Pro Val Gly Ala Ala Ala Val Pro Leu
 340 345 350
 Leu Ser Ala Gly Gly Ala Ala Pro Pro His Pro Gly Pro Asp Ala Ala
 355 360 365
 Val Phe Arg Ser Ser Leu Gly Ser Leu Leu Tyr Trp Pro Gly Val Arg
 370 375 380
 Ala Leu Leu Gly Arg Asp Cys Arg Val Ala Ala Arg Tyr Ala Gly Arg
 385 390 395 400
 Met Thr Tyr Ile Ala Thr Gly Ala Leu Leu Ala Arg Phe Asn Pro Gly
 405 410 415
 Ala Val Lys Cys Val Leu Pro Arg Glu Ala Ala Phe Ala Gly Arg Val
 420 425 430
 Leu Asp Val Leu Ala Val Leu Ala Glu Gln Thr Val Gln Trp Leu Ser
 435 440 445
 Val Val Val Gly Ala Arg Leu His Pro His Ser Ala His Pro Ala Phe
 450 455 460
 Ala Asp Val Glu Gln Glu Ala Leu Phe Arg Ala Leu Pro Leu Gly Ser
 465 470 475 480
 Pro Gly Val Val Ala Ala Glu His Glu Ala Leu Gly Asp Thr Ala Ala
 485 490 495
 Arg Arg Leu Leu Ala Thr Ser Gly Leu Asn Ala Val Leu Gly Ala Ala
 500 505 510
 Val Tyr Ala Leu His Thr Ala Leu Ala Thr Val Thr Leu Lys Tyr Ala
 515 520 525

Leu Ala Cys Gly Asp Ala Arg Arg Arg Arg Asp Asp Ala Ala Ala Ala
 530 535 540
 Arg Ala Val Leu Ala Thr Gly Leu Ile Leu Gln Arg Leu Leu Gly Leu
 545 550 555 560
 Ala Asp Thr Val Val Ala Cys Val Ala Leu Ala Ala Phe Asp Gly Gly
 565 570 575
 Ser Thr Ala Pro Glu Val Gly Thr Tyr Thr Pro Leu Arg Tyr Ala Cys
 580 585 590
 Val Leu Arg Ala Thr Gln Pro Leu Tyr Ala Arg Thr Thr Pro Ala Lys
 595 600 605
 Phe Trp Ala Asp Val Arg Ala Ala Ala Glu His Val Asp Leu Arg Pro
 610 615 620
 Ala Ser Ser Ala Pro Arg Ala Pro Val Ser Gly Thr Ala Asp Pro Ala
 625 630 635 640
 Phe Leu Leu Glu Asp Leu Ala Ala Phe Pro Pro Ala Pro Leu Asn Ser
 645 650 655
 Glu Ser Val Leu Gly Pro Arg Val Arg Val Val Asp Ile Met Ala Gln
 660 665 670
 Phe Arg Lys Leu Leu Met Gly Asp Glu Glu Thr Ala Ala Leu Arg Ala
 675 680 685
 His Val Ser Gly Arg Arg Ala Thr Gly Leu Gly Gly Pro Pro Arg Pro
 690 695 700

<210> 95
 <211> 1381
 <212> PRT
 <213> Homo sapiens

<400> 95
 Met Gln His His His His His His Ala Ala Pro Ala Arg Asp Pro Pro
 5 10 15
 Gly Tyr Arg Tyr Ala Ala Ala Met Val Pro Thr Gly Ser Ile Leu Ser
 20 25 30
 Thr Ile Glu Val Ala Ser His Arg Arg Leu Phe Asp Phe Phe Ala Arg
 35 40 45
 Val Arg Ser Asp Glu Asn Ser Leu Tyr Asp Val Glu Phe Asp Ala Leu
 50 55 60
 Leu Gly Ser Tyr Cys Asn Thr Leu Ser Leu Val Arg Phe Leu Glu Leu
 65 70 75 80
 Gly Leu Ser Val Ala Cys Val Cys Thr Lys Phe Pro Glu Leu Ala Tyr
 85 90 95

Met Asn Glu Gly Arg Val Gln Phe Glu Val His Gln Pro Leu Ile Ala
 100 105 110
 Arg Asp Gly Pro His Pro Val Glu Gln Pro Val His Asn Tyr Met Thr
 115 120 125
 Lys Val Ile Asp Arg Arg Ala Leu Asn Ala Ala Phe Ser Leu Ala Thr
 130 135 140
 Glu Ala Ile Ala Leu Leu Thr Gly Glu Ala Leu Asp Gly Thr Gly Ile
 145 150 155 160
 Ser Leu His Arg Gln Leu Arg Ala Ile Gln Gln Leu Ala Arg Asn Val
 165 170 175
 Gln Ala Val Leu Gly Ala Phe Glu Arg Gly Thr Ala Asp Gln Met Leu
 180 185 190
 His Val Leu Leu Glu Lys Ala Pro Pro Leu Ala Leu Leu Leu Pro Met
 195 200 205
 Gln Arg Tyr Leu Asp Asn Gly Arg Leu Ala Thr Arg Val Ala Arg Ala
 210 215 220
 Thr Leu Val Ala Glu Leu Lys Arg Ser Phe Cys Asp Thr Ser Phe Phe
 225 230 235 240
 Leu Gly Lys Ala Gly His Arg Arg Glu Ala Ile Glu Ala Trp Leu Val
 245 250 255
 Asp Leu Thr Thr Ala Thr Gln Pro Ser Val Ala Val Pro Arg Leu Thr
 260 265 270
 His Ala Asp Thr Arg Gly Arg Pro Val Asp Gly Val Leu Val Thr Thr
 275 280 285
 Ala Ala Ile Lys Gln Arg Leu Leu Gln Ser Phe Leu Lys Val Glu Asp
 290 295 300
 Thr Glu Ala Asp Val Pro Val Thr Tyr Gly Glu Met Val Leu Asn Gly
 305 310 315 320
 Ala Asn Leu Val Thr Ala Leu Val Met Gly Lys Ala Val Arg Ser Leu
 325 330 335
 Asp Asp Val Gly Arg His Leu Leu Glu Met Gln Glu Glu Gln Leu Glu
 340 345 350
 Ala Asn Arg Glu Thr Leu Asp Glu Leu Glu Ser Ala Pro Gln Thr Thr
 355 360 365
 Arg Val Arg Ala Asp Leu Val Ala Ile Gly Asp Arg Leu Val Phe Leu
 370 375 380
 Glu Ala Leu Glu Lys Arg Ile Tyr Ala Ala Thr Asn Val Pro Tyr Pro
 385 390 395 400
 Leu Val Gly Ala Met Asp Leu Thr Phe Val Leu Pro Leu Gly Leu Phe

405										410					415															
Asn	Pro	Ala	Met	Glu	Arg	Phe	Ala	Ala	His	Ala	Gly	Asp	Leu	Val	Pro															
			420					425					430																	
Ala	Pro	Gly	His	Pro	Glu	Pro	Arg	Ala	Phe	Pro	Pro	Arg	Gln	Leu	Phe															
		435					440					445																		
Phe	Trp	Gly	Lys	Asp	His	Gln	Val	Leu	Arg	Leu	Ser	Met	Glu	Asn	Ala															
	450					455					460																			
Val	Gly	Thr	Val	Cys	His	Pro	Ser	Leu	Met	Asn	Ile	Asp	Ala	Ala	Val															
465					470					475																				
Gly	Gly	Val	Asn	His	Ala	Pro	Val	Glu	Ala	Ala	Asn	Pro	Tyr	Gly	Ala															
				485					490					495																
Tyr	Val	Ala	Ala	Pro	Ala	Gly	Pro	Gly	Ala	Asp	Met	Gln	Gln	Arg	Phe															
		500						505					510																	
Leu	Asn	Ala	Trp	Arg	Gln	Arg	Leu	Ala	His	Gly	Arg	Val	Arg	Trp	Val															
		515					520					525																		
Ala	Glu	Cys	Gln	Met	Thr	Ala	Glu	Gln	Phe	Met	Gln	Pro	Asp	Asn	Ala															
	530					535					540																			
Asn	Leu	Ala	Leu	Glu	Leu	His	Pro	Ala	Phe	Asp	Phe	Phe	Ala	Gly	Val															
545					550					555				560																
Ala	Asp	Val	Glu	Leu	Pro	Gly	Gly	Glu	Val	Pro	Pro	Ala	Gly	Pro	Gly															
				565					570					575																
Ala	Ile	Gln	Ala	Thr	Trp	Arg	Val	Val	Asn	Gly	Asn	Leu	Pro	Leu	Ala															
		580						585					590																	
Leu	Cys	Pro	Val	Ala	Phe	Arg	Asp	Ala	Arg	Gly	Leu	Glu	Leu	Gly	Val															
		595					600					605																		
Gly	Arg	His	Ala	Met	Ala	Pro	Ala	Thr	Ile	Ala	Ala	Val	Arg	Gly	Ala															
	610					615					620																			

Leu Gly Gly Gln Ala Gln Ala Glu Leu Asn His Leu Met Arg Asp Pro
 725 730 735
 Ala Leu Leu Pro Pro Leu Val Trp Asp Cys Asp Gly Leu Met Arg His
 740 745 750
 Ala Ala Leu Asp Arg His Arg Asp Cys Arg Ile Asp Ala Gly Gly His
 755 760 765
 Glu Pro Val Tyr Ala Ala Ala Cys Asn Val Ala Thr Ala Asp Phe Asn
 770 775 780
 Arg Asn Asp Gly Arg Leu Leu His Asn Thr Gln Ala Arg Ala Val Asp
 785 790 795 800
 Ala Ala Asp Asp Arg Pro His Arg Pro Ala Asp Trp Thr Val His His
 805 810 815
 Lys Ile Tyr Tyr Tyr Val Leu Val Pro Ala Phe Ser Arg Gly Arg Cys
 820 825 830
 Cys Thr Ala Gly Val Arg Phe Asp Arg Val Tyr Ala Thr Leu Gln Asn
 835 840 845
 Met Val Val Pro Glu Ile Ala Pro Gly Glu Glu Cys Pro Ser Asp Pro
 850 855 860
 Val Thr Asp Pro Ala His Pro Leu His Pro Ala Asn Leu Val Ala Asn
 865 870 875 880
 Thr Val Asn Ala Met Phe His Asn Gly Arg Val Val Val Asp Gly Pro
 885 890 895
 Ala Met Leu Thr Leu Gln Val Leu Ala His Asn Met Ala Glu Arg Thr
 900 905 910
 Thr Ala Leu Leu Cys Ser Ala Ala Pro Asp Ala Gly Ala Asn Thr Ala
 915 920 925
 Ser Thr Ala Asn Met Arg Ile Phe Asp Gly Ala Leu His Ala Gly Val
 930 935 940
 Leu Leu Met Ala Pro Gln His Leu Asp His Thr Ile Gln Asn Gly Glu
 945 950 955 960
 Tyr Phe Tyr Val Leu Pro Val His Ala Leu Phe Ala Gly Ala Asp His
 965 970 975
 Val Ala Asn Ala Pro Asn Phe Pro Pro Ala Leu Arg Asp Leu Ala Arg
 980 985 990
 His Val Pro Leu Val Pro Pro Ala Leu Gly Ala Asn Tyr Phe Ser Ser
 995 1000 1005
 Ile Arg Gln Pro Val Val Gln His Ala Arg Glu Ser Ala Ala Gly Glu
 1010 1015 1020

Asn Ala Leu Thr Tyr Ala Leu Met Ala Gly Tyr Phe Lys Met Ser Pro
 1025 1030 1035 1040
 Val Ala Leu Tyr His Gln Leu Lys Thr Gly Leu His Pro Gly Phe Gly
 1045 1050 1055
 Phe Thr Val Val Arg Gln Asp Arg Phe Val Thr Glu Asn Val Leu Phe
 1060 1065 1070
 Ser Glu Arg Ala Ser Glu Ala Tyr Phe Leu Gly Gln Leu Gln Val Ala
 1075 1080 1085
 Arg His Glu Thr Gly Gly Gly Val Ser Phe Thr Leu Thr Gln Pro Arg
 1090 1095 1100
 Gly Asn Val Asp Leu Gly Val Gly Tyr Thr Ala Val Ala Ala Thr Ala
 1105 1110 1115 1120
 Thr Val Arg Asn Pro Val Thr Asp Met Gly Asn Leu Pro Gln Asn Phe
 1125 1130 1135
 Tyr Leu Gly Arg Gly Ala Pro Pro Leu Leu Asp Asn Ala Ala Ala Val
 1140 1145 1150
 Tyr Leu Arg Asn Ala Val Val Ala Gly Asn Arg Leu Gly Pro Ala Gln
 1155 1160 1165
 Pro Leu Pro Val Phe Gly Cys Ala Gln Val Pro Arg Arg Ala Gly Met
 1170 1175 1180
 Asp His Gly Gln Asp Ala Val Cys Glu Phe Ile Ala Thr Pro Val Ala
 1185 1190 1195 1200
 Thr Asp Ile Asn Tyr Phe Arg Arg Pro Cys Asn Pro Arg Gly Arg Ala
 1205 1210 1215
 Ala Gly Gly Val Tyr Ala Gly Asp Lys Glu Gly Asp Val Ile Ala Leu
 1220 1225 1230
 Met Tyr Asp His Gly Gln Ser Asp Pro Ala Arg Pro Phe Ala Ala Thr
 1235 1240 1245
 Ala Asn Pro Trp Ala Ser Gln Arg Phe Ser Tyr Gly Asp Leu Leu Tyr
 1250 1255 1260
 Asn Gly Ala Tyr His Leu Asn Gly Ala Ser Pro Val Leu Ser Pro Cys
 1265 1270 1275 1280
 Phe Lys Phe Phe Thr Ala Ala Asp Ile Thr Ala Lys His Arg Cys Leu
 1285 1290 1295
 Glu Arg Leu Ile Val Glu Thr Gly Ser Ala Val Ser Thr Ala Thr Ala
 1300 1305 1310
 Ala Ser Asp Val Gln Phe Lys Arg Pro Pro Gly Cys Arg Glu Leu Val
 1315 1320 1325
 Glu Asp Pro Cys Gly Leu Phe Gln Glu Ala Tyr Pro Ile Thr Cys Ala

1330 1335 1340
 Ser Asp Pro Ala Leu Leu Arg Ser Ala Arg Asp Gly Glu Ala His Ala
 1345 1350 1355 1360
 Arg Glu Thr His Phe Thr Gln Tyr Leu Ile Tyr Asp Ala Ser Pro Leu
 1365 1370 1375
 Lys Gly Leu Ser Leu
 1380

 <210> 96
 <211> 377
 <212> PRT
 <213> Homo sapiens

 <400> 96
 Met Gln His His His His His His His Ser Gln Trp Gly Pro Arg Ala
 5 10 15
 Ile Leu Val Gln Thr Asp Ser Thr Asn Arg Asn Ala Asp Gly Asp Trp
 20 - 25 30
 Gln Ala Ala Val Ala Ile Arg Gly Gly Gly Val Val Gln Leu Asn Met
 35 40 45
 Val Asn Lys Arg Ala Val Asp Phe Thr Pro Ala Glu Cys Gly Asp Ser
 50 55 60
 Glu Trp Ala Val Gly Arg Val Ser Leu Gly Leu Arg Met Ala Met Pro
 65 70 75 80
 Arg Asp Phe Cys Ala Ile Ile His Ala Pro Ala Val Ser Gly Pro Gly
 85 90 95
 Pro His Val Met Leu Gly Leu Val Asp Ser Gly Tyr Arg Gly Thr Val
 100 105 110
 Leu Ala Val Val Val Ala Pro Asn Gly Thr Arg Gly Phe Ala Pro Gly
 115 120 125
 Ala Leu Arg Val Asp Val Thr Phe Leu Asp Ile Arg Ala Thr Pro Pro
 130 135 140
 Thr Leu Thr Glu Pro Ser Ser Leu His Arg Phe Pro Gln Leu Ala Pro
 145 150 155 160
 Ser Pro Leu Ala Gly Leu Arg Glu Asp Pro Trp Leu Asp Gly Ala Leu
 165 170 175
 Ala Thr Ala Gly Gly Ala Val Ala Leu Pro Ala Arg Arg Arg Gly Gly
 180 185 190
 Ser Leu Val Tyr Ala Gly Glu Leu Thr Gln Val Thr Thr Glu His Gly
 195 200 205
 Asp Cys Val His Glu Ala Pro Ala Phe Leu Pro Lys Arg Glu Glu Asp

```
<210> 97
<211> 308
<212> PRT
<213> Homo sapiens
```

```

<400> 97
Met Gln His His His His His His His Thr Ser Arg Arg Ser Val Lys
                    5                      10                      15

Ser Cys Pro Arg Glu Ala Pro Arg Gly Thr His Glu Glu Leu Tyr Tyr
                20                      25                      30

Gly Pro Val Ser Pro Ala Asp Pro Glu Ser Pro Arg Asp Asp Phe Arg
                35                      40                      45

Arg Gly Ala Gly Pro Met Arg Ala Arg Pro Arg Gly Glu Val Arg Phe
                50                      55                      60

Leu His Tyr Asp Glu Ala Gly Tyr Ala Leu Tyr Arg Asp Ser Ser Ser
    65                      70                      75                      80

Asp Asp Asp Glu Ser Arg Asp Thr Ala Arg Pro Arg Arg Ser Ala Ser
                85                      90                      95

Val Ala Gly Ser His Gly Pro Gly Pro Ala Arg Ala Pro Pro Pro Pro

```

100					105					110						
Gly	Gly	Pro	Val	Gly	Ala	Gly	Gly	Arg	Ser	His	Ala	Pro	Pro	Ala	Arg	
115					120					125						
Thr	Pro	Lys	Met	Thr	Arg	Gly	Ala	Pro	Lys	Ala	Pro	Ala	Thr	Pro	Ala	
130					135					140						
Thr	Asp	Pro	Ala	Arg	Gly	Arg	Arg	Pro	Ala	Gln	Ala	Asp	Ser	Ala	Val	
145					150					155					160	
Leu	Leu	Asp	Ala	Pro	Ala	Pro	Thr	Ala	Ser	Gly	Arg	Thr	Lys	Thr	Pro	
165					170					175						
Ala	Gln	Gly	Leu	Ala	Lys	Lys	Leu	His	Phe	Ser	Thr	Ala	Pro	Pro	Ser	
180					185					190						
Pro	Thr	Ala	Pro	Trp	Thr	Pro	Arg	Val	Ala	Gly	Phe	Asn	Lys	Arg	Val	
195					200					205						
Phe	Cys	Ala	Ala	Val	Gly	Arg	Leu	Ala	Ala	Thr	His	Ala	Arg	Leu	Ala	
210					215					220						
Ala	Val	Gln	Leu	Trp	Asp	Met	Ser	Arg	Pro	His	Thr	Asp	Glu	Asp	Leu	
225					230					235					240	
Asn	Glu	Leu	Leu	Asp	Leu	Thr	Thr	Ile	Arg	Val	Thr	Val	Cys	Glu	Gly	
245					250					255						
Lys	Asn	Leu	Leu	Gln	Arg	Ala	Asn	Glu	Leu	Val	Asn	Pro	Asp	Ala	Ala	
260					265					270						
Gln	Asp	Val	Asp	Ala	Thr	Ala	Ala	Ala	Arg	Gly	Arg	Pro	Ala	Gly	Arg	
275					280					285						
Ala	Ala	Ala	Thr	Ala	Arg	Ala	Pro	Ala	Arg	Ser	Ala	Ser	Arg	Pro	Arg	
290					295					300						
Arg	Pro	Leu	Glu													
305																

<210> 98
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 98
 gagctcagct atgccaccac c

<210> 99
 <211> 33
 <212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 99

cggcgaattc attagtagag gcggtggaaa aag

33

<210> 100

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 100

cacgcgcgcg caccacagc ggac

24

<210> 101

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 101

cggcgaattc attagtagag gcggtggaaa aag

33

<210> 102

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 102

cacacctctc gccgtccgt caagtc

26

<210> 103

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 103

cataagaatt cactactcga gggggcggcg gggacg

36

<210> 104

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 104

cacagtcagt gggggcccag ggcgatcc

28

<210> 105

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 105

cctagaattc actagatgcc agtggagcca aaccc

35

<210> 106

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 106

gccgctcctg cccgcgaccc ccc

23

<210> 107

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 107

ccagaattca ttacagagac aggcccttta gc

32

<210> 108

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 108

cactccgtgg cgcgggcatg ccg

23

<210> 109

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 109

ccgttagaat tcactatggg cgtggcgggc c 31

<210> 110
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 110
cactccgtgc gcgggcatgc cg 22

<210> 111
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 111
catagaattc atcacgcgcg ggaggggctg gtttttgc 38

<210> 112
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 112
gacacggtgg tcgcgtgcgt ggc 23

<210> 113
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 113
ccgttagaat tcactatggg cgtggcgggc c 31

<210> 114
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> PCR primer

<400> 114
cactccgtgc gcgggcatgc cg 22

<210> 115

<211> 28
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 115
 cgtatgaatt catcagaccc acccgttg

28

<210> 116
 <211> 22
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 116
 gtgctggcga cggggctcat cc

22

<210> 117
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 117
 ccgtagaat tcactatggg cgtggcgggc c

31

<210> 118
 <211> 783
 <212> DNA
 <213> HSV-2

<400> 118
 atggctcgcg gggccggggt ggtgtttttt gttggagttt gggtcgtatc gtgcctggcg 60
 gcagcaccca gaacgtcctg gaaacgggta acctcgggag aggacgtggt gttgcttccg 120
 gcgcccgcgg aacgcacccg ggcccacaaa ctactatggg ccgcggaacc cctggatgcc 180
 tgcggtcccc tgcgcccgtc gtgggtggcg ctgtggcccc cccgacgggt gctcgagacg 240
 gtcgtggatg cggcgtgcat gcgcgccccg gaaccgctcg ccatagcata cagtcccccg 300
 ttccccgcgg gcgacgaggg actgtattcg gagttggcgt ggcgcgatcg cgtagccgtg 360
 gtcaacgaga gtctggatcat ctacggggcc ctggagacgg acagcgggtc gtacaccctg 420
 tccgtggtcg gcctaagcga cgaggcgcgc caagtggcgt cgggtggttct ggtcgtggag 480
 cccgcccctg tgccgacccc gacccccgac gactacgacg aagaagacga cgcgggcgtg 540
 agcgaacgca cgcgggtcag cgttcccccc ccaaccccc ccccgctcgc cccccgtcgc 600
 cccccgaacg caccctcgtg ttatccccga ggtgtccac gtgcgcgggg taacgggtcca 660
 tatggagacc cgggagcca ttctgtttgc ccccggggag acgtttggga cgaacgtctc 720
 catccacgcc attgccacg acgacgggtc gtacgccatg gacgtcgtct ggatgcggtt 780
 tga 783

<210> 119
 <211> 1638
 <212> DNA
 <213> HSV-2

<400> 119
 atggctcgcg gggccgggtt ggtgtttttt gttggagttt gggtcgtatc gtgcctggcg 60
 gcagcaccga gaacgtcctg gaaacgggta acctcgggcg aggacgtggt gttgcttccg 120
 gcgcccgcgg aacgcacccg ggcccacaaa ctactgtggg ccgcggaacc cctggatgcc 180
 tgcggtcccc tgcgcccgtc gtgggtggcg ctgtggcccc ccgcacgggt gctcgagacg 240
 gtcgtggatg cggcgtgcat gcgcgccccg gaaccgctcg ccatagcata cagtcccccg 300
 ttccccgcgg gcgacgaggg actgtattcg gagttggcgt ggcgcgatcg cgtagccgtg 360
 gtcaacgaga gtctggtcat ctacggggcc ctggagacgg acagcgggtc gtacaccctg 420
 tccgtggtcg gcctaagcga cgaggcgcg caagtggcgt cgggtggttct ggtcgtggag 480
 cccgccccctg tgccgacccc gacccccgac gactacgacg aagaagacga cgcgggcgtg 540
 acgaacgcac gccgggtcagc gttccccccc caaccccccc cccgtcgtcc ccccgctcgc 600
 cccccgacgc accctcgtgt tatccccgag gtgtcccacg tgcgcggggg aacggtccat 660
 atggagaccc tggaggccat tctgtttgcc cccggggaga cgtttgggac gaacgtctcc 720
 atccacgcca ttgccacga cgacggtccg tacgccatgg acgtcgtctg gatgcggttt 780
 gacgtgccgt cctcgtgcgc cgatatgcgg atctacgaag cttgtctgta tcaccgcag 840
 cttccagagt gtctatctcc ggccgacgcy ccgtgcgcgc taagttcctg ggcgtaccgc 900
 ctggcgggtcc gcagctacgc cggctgttcc aggactacgc ccccgccgcg atgttttgcc 960
 gaggtcgcga tggaaacggg cccgggggtg gcgtggctgg cctccaccgt caatctggaa 1020
 ttccagcacg cctcccccca gcacgcggc ctctacctgt gcgtgggtga cgtggacgat 1080
 catatccacg cctggggcca catgaccatc agcacgcgg cgcagtaccg gaacgcggtg 1140
 gtggaacagc acctccccca gcgccagccc gagcccgtcg agcccaccg cccgcacgtg 1200
 agagcccccc atccgcgcgc ctccgcgcgc ggcccgtgc gcctcggggc ggtgctgggg 1260
 gcggccctgt tgcgtggcgc cctcgggctg tccgcgtggg cgtgcatgac ctgctggcgc 1320
 aggcgctcct ggcggggcgg taaaagccgg gcctcggcga cgggccccac ttacattcgc 1380
 gtggcggaca gcgagctgta cgcggactgg agttcggaca gcgaggggga gcgcgacggg 1440
 tccctgtggc aggaccctcc ggagagaccc gactctccct ccacaaatgg atccggcttt 1500
 gagatcttat caccaacggc tccgtctgta tacccccata gcgaggggcy taaatctcgc 1560
 cgcgcgtca ccacctttg ttcggaagc cccggccgct gtcactccca ggcctcctat 1620
 ccgtccgtcc tctggtaa 1638

<210> 120
 <211> 260
 <212> PRT
 <213> HSV-2

<400> 120
 Met Ala Arg Gly Ala Gly Leu Val Phe Phe Val Gly Val Trp Val Val
 5 10 15
 Ser Cys Leu Ala Ala Ala Pro Arg Thr Ser Trp Lys Arg Val Thr Ser
 20 25 30
 Gly Glu Asp Val Val Leu Leu Pro Ala Pro Ala Glu Arg Thr Arg Ala
 35 40 45
 His Lys Leu Leu Trp Ala Ala Glu Pro Leu Asp Ala Cys Gly Pro Leu
 50 55 60
 Arg Pro Ser Trp Val Ala Leu Trp Pro Pro Arg Arg Val Leu Glu Thr
 65 70 75 80
 Val Val Asp Ala Ala Cys Met Arg Ala Pro Glu Pro Leu Ala Ile Ala
 85 90 95
 Tyr Ser Pro Pro Phe Pro Ala Gly Asp Glu Gly Leu Tyr Ser Glu Leu
 100 105 110

Ala Trp Arg Asp Arg Val Ala Val Val Asn Glu Ser Leu Val Ile Tyr
 115 120 125

Gly Ala Leu Glu Thr Asp Ser Gly Leu Tyr Thr Leu Ser Val Val Gly
 130 135 140

Leu Ser Asp Glu Ala Arg Gln Val Ala Ser Val Val Leu Val Val Glu
 145 150 155 160

Pro Ala Pro Val Pro Thr Pro Thr Pro Asp Asp Tyr Asp Glu Glu Asp
 165 170 175

Asp Ala Gly Val Ser Glu Arg Thr Pro Val Ser Val Pro Pro Pro Thr
 180 185 190

Pro Pro Pro Ser Ser Pro Arg Arg Pro Pro Asp Ala Pro Ser Cys Tyr
 195 200 205

Pro Arg Gly Val Pro Arg Ala Arg Gly Asn Gly Pro Tyr Gly Asp Pro
 210 215 220

Gly Gly His Ser Val Cys Pro Arg Gly Asp Val Trp Asp Glu Arg Leu
 225 230 235 240

His Pro Arg His Cys Pro Arg Arg Arg Ser Val Arg His Gly Arg Arg
 245 250 255

Leu Asp Ala Val
 260

<210> 121
 <211> 545
 <212> PRT
 <213> HSV-2

<400> 121

Met Ala Arg Gly Ala Gly Leu Val Phe Phe Val Gly Val Trp Val Val
 5 10 15

Ser Cys Leu Ala Ala Ala Pro Arg Thr Ser Trp Lys Arg Val Thr Ser
 20 25 30

Gly Glu Asp Val Val Leu Leu Pro Ala Pro Ala Glu Arg Thr Arg Ala
 35 40 45

His Lys Leu Leu Trp Ala Ala Glu Pro Leu Asp Ala Cys Gly Pro Leu
 50 55 60

Arg Pro Ser Trp Val Ala Leu Trp Pro Pro Arg Arg Val Leu Glu Thr
 65 70 75 80

Val Val Asp Ala Ala Cys Met Arg Ala Pro Glu Pro Leu Ala Ile Ala
 85 90 95

Tyr Ser Pro Pro Phe Pro Ala Gly Asp Glu Gly Leu Tyr Ser Glu Leu
 100 105 110

Ala Trp Arg Asp Arg Val Ala Val Val Asn Glu Ser Leu Val Ile Tyr
 115 120 125
 Gly Ala Leu Glu Thr Asp Ser Gly Leu Tyr Thr Leu Ser Val Val Gly
 130 135 140
 Leu Ser Asp Glu Ala Arg Gln Val Ala Ser Val Val Leu Val Val Glu
 145 150 155 160
 Pro Ala Pro Val Pro Thr Pro Thr Pro Asp Asp Tyr Asp Glu Glu Asp
 165 170 175
 Asp Ala Gly Val Thr Asn Ala Arg Arg Ser Ala Phe Pro Pro Gln Pro
 180 185 190
 Pro Pro Arg Arg Pro Pro Val Ala Pro Pro Thr His Pro Arg Val Ile
 195 200 205
 Pro Glu Val Ser His Val Arg Gly Val Thr Val His Met Glu Thr Leu
 210 215 220
 Glu Ala Ile Leu Phe Ala Pro Gly Glu Thr Phe Gly Thr Asn Val Ser
 225 230 235 240
 Ile His Ala Ile Ala His Asp Asp Gly Pro Tyr Ala Met Asp Val Val
 245 250 255
 Trp Met Arg Phe Asp Val Pro Ser Ser Cys Ala Asp Met Arg Ile Tyr
 260 265 270
 Glu Ala Cys Leu Tyr His Pro Gln Leu Pro Glu Cys Leu Ser Pro Ala
 275 280 285
 Asp Ala Pro Cys Ala Val Ser Ser Trp Ala Tyr Arg Leu Ala Val Arg
 290 295 300
 Ser Tyr Ala Gly Cys Ser Arg Thr Thr Pro Pro Arg Cys Phe Ala
 305 310 315 320
 Glu Ala Arg Met Glu Pro Val Pro Gly Leu Ala Trp Leu Ala Ser Thr
 325 330 335
 Val Asn Leu Glu Phe Gln His Ala Ser Pro Gln His Ala Gly Leu Tyr
 340 345 350
 Leu Cys Val Val Tyr Val Asp Asp His Ile His Ala Trp Gly His Met
 355 360 365
 Thr Ile Ser Thr Ala Ala Gln Tyr Arg Asn Ala Val Val Glu Gln His
 370 375 380
 Leu Pro Gln Arg Gln Pro Glu Pro Val Glu Pro Thr Arg Pro His Val
 385 390 395 400
 Arg Ala Pro His Pro Ala Pro Ser Ala Arg Gly Pro Leu Arg Leu Gly
 405 410 415
 Ala Val Leu Gly Ala Ala Leu Leu Leu Ala Ala Leu Gly Leu Ser Ala

420	425	430
Trp Ala Cys Met Thr Cys Trp Arg Arg Arg Ser Trp Arg Ala Val Lys		
435	440	445
Ser Arg Ala Ser Ala Thr Gly Pro Thr Tyr Ile Arg Val Ala Asp Ser		
450	455	460
Glu Leu Tyr Ala Asp Trp Ser Ser Asp Ser Glu Gly Glu Arg Asp Gly		
465	470	475
Ser Leu Trp Gln Asp Pro Pro Glu Arg Pro Asp Ser Pro Ser Thr Asn		
485	490	495
Gly Ser Gly Phe Glu Ile Leu Ser Pro Thr Ala Pro Ser Val Tyr Pro		
500	505	510
His Ser Glu Gly Arg Lys Ser Arg Arg Pro Leu Thr Thr Phe Gly Ser		
515	520	525
Gly Ser Pro Gly Arg Arg His Ser Gln Ala Ser Tyr Pro Ser Val Leu		
530	535	540
Trp		
545		

<210> 122
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic peptide

<400> 122
 Val Gly Ala Ala Ala Val Pro Leu Leu Ser Ala Gly Gly Ala Ala
 1 5 10 15

<210> 123
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic peptide

<400> 123
 Pro His Pro Gly Pro Asp Ala Ala Val Phe Arg Ser Ser Leu Gly
 1 5 10 15

<210> 124
 <211> 11
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 124

Met	Thr	Tyr	Ile	Ala	Thr	Gly	Ala	Leu	Leu	Ala
1				5				10		

<210> 125

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 125

Glu	Ala	Ala	Phe	Ala	Gly	Arg	Val	Leu	Asp	Val	Leu	Ala	Val	Leu
1				5				10					15	

<210> 126

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 126

Ala	Arg	Leu	His	Pro	His	Ser	Ala	His	Pro	Ala	Phe	Ala	Asp	Val
1				5				10					15	

<210> 127

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 127

Ser	Pro	Asn	Thr	Asp	Val	Arg	Met	Tyr	Ser	Gly	Lys	Arg	Asn	Gly
				5				10					15	

<210> 128

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 128

Tyr	Leu	Ala	Ala	Pro	Thr	Gly	Ile	Pro	Pro	Ala	Phe	Phe	Pro	Ile
				5				10					15	

<210> 129
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 129
Gly Val Ala Ala Ala Thr Pro Arg Pro Asp Pro Glu Asp Gly Ala
 5 10 15

<210> 130
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 130
Glu Glu Ile Pro Trp Val Arg Val Tyr Glu Asn Ile Cys Pro Arg
 5 10 15

<210> 131
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 131
Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala Glu Asn Pro
 5 10 15

<210> 132
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 132
Pro Asp Ser Pro Tyr Ile Glu Ala Glu Asn Pro Leu Tyr Asp Trp
 5 10 15

<210> 133
<211> 15
<212> PRT
<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 133

Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro
5 10 15

<210> 134

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 134

Ala Val Pro Leu Leu Ser Ala Gly Gly Ala Ala Pro Pro His Pro
5 10 15

<210> 135

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 135

Glu Leu Tyr Tyr Gly Pro Val Ser Pro Ala Asp Pro Glu Ser Pro
5 10 15

<210> 136

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 136

Pro Met Arg Ala Arg Pro Arg Gly Glu Val Arg Phe Leu His Tyr
5 10 15

<210> 137

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 137

Arg Pro Arg Gly Glu Val Arg Phe Leu His Tyr Asp Glu Ala Gly

5

10

15

<210> 138
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic peptide

<400> 138
 Val Ala Gly Phe Asn Lys Arg Val Phe Cys Ala Ala Val Gly Arg
 5 10 15

<210> 139
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic peptide

<400> 139
 Ala Ile Asp Tyr Val His Cys Glu Gly Ile Ile His Arg Asp Ile
 5 10 15

<210> 140
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic peptide

<400> 140
 Ala Phe Pro Val Ala Leu His Ala Val Asp Ala Pro Ser Gln Phe
 5 10 15

<210> 141
 <211> 1808
 <212> DNA
 <213> HSV-2

<400> 141
 atggacggg aggcacttcg ggccatcagc cgcgggtgca agcccccttc gaccctggca 60
 aaactggtga cggggtggg attcgcgatc caggagcgc tcatcccggt gtcggagggg 120
 tgtgtctttg atagcagcca cccgaactac cctcatcggg taatcgtcaa ggcgggggtg 180
 tacgccagca cgaaccacga ggcgcggctg ctgagacgcc tgaaccaccc cgcgatccta 240
 cccctccttg acctgcacgt cgtttctggg gtcacgtgtc tggtcctccc caagtatcac 300
 tgcgacctgt atacctatct gagcaagcgc ccgtctccgt tgggccacct acagataacc 360
 gcggtctccc ggcagctctt gagcgccatc gactacgtcc actgcgaagg catcatccac 420
 cgcgatatta agaccgagaa catcctcatc aacacccccg agaacatctg tctgggggac 480
 tttggggcgg cgtgctttgt gcgcgggtgt cgatcgagcc ccttccatta cgggatcgca 540

<210>	142
<211>	248
<212>	PRT
<213>	HSV-2

```

<400> 142
Met His Ala Ile Ala Pro Arg Leu Leu Leu Leu Phe Val Leu Ser Gly
          5                      10                      15

Leu Pro Gly Thr Arg Gly Gly Ser Gly Val Pro Gly Pro Ile Asn Pro
          20                      25                      30

Pro Asn Asn Asp Val Val Phe Pro Gly Gly Ser Pro Val Ala Gln Tyr
          35                      40                      45

Cys Tyr Ala Tyr Pro Arg Leu Asp Asp Pro Gly Pro Leu Gly Ser Ala
          50                      55                      60

Asp Ala Gly Arg Gln Asp Leu Pro Arg Arg Val Val Arg His Glu Pro
          65                      70                      75                      80

Leu Gly Arg Ser Phe Leu Thr Gly Gly Leu Val Leu Leu Ala Pro Pro
          85                      90                      95

Val Arg Gly Phe Gly Ala Pro Asn Ala Thr Tyr Ala Ala Arg Val Thr
          100                     105                     110

Tyr Tyr Arg Leu Thr Arg Ala Cys Arg Gln Pro Ile Leu Leu Arg Gln
          115                     120                     125

Tyr Gly Gly Cys Arg Gly Gly Glu Pro Pro Ser Pro Lys Thr Cys Gly
          130                     135                     140

Ser Tyr Thr Tyr Thr Tyr Gln Gly Gly Gly Pro Pro Thr Arg Tyr Ala
          145                     150                     155                     160

```

Leu Val Asn Ala Ser Leu Leu Val Pro Ile Trp Asp Arg Ala Ala Glu
 165 170 175
 Thr Phe Glu Tyr Gln Ile Glu Leu Gly Gly Glu Leu His Val Gly Leu
 180 185 190
 Leu Trp Val Glu Val Gly Gly Glu Gly Pro Gly Pro Thr Ala Pro Pro
 195 200 205
 Gln Ala Ala Arg Ala Glu Gly Gly Pro Cys Val Pro Pro Val Pro Ala
 210 215 220
 Gly Arg Pro Trp Arg Ser Val Pro Pro Val Trp Tyr Ser Ala Pro Asn
 225 230 235 240
 Pro Gly Phe Arg Gly Leu Arg Phe
 245

<210> 143
 <211> 699
 <212> PRT
 <213> HSV-2

<400> 143

Met His Ala Ile Ala Pro Arg Leu Leu Leu Leu Phe Val Leu Ser Gly
 5 10 15
 Leu Pro Gly Thr Arg Gly Gly Ser Gly Val Pro Gly Pro Ile Asn Pro
 20 25 30
 Pro Asn Ser Asp Val Val Phe Pro Gly Gly Ser Pro Val Ala Gln Tyr
 35 40 45
 Cys Tyr Ala Tyr Pro Arg Leu Asp Asp Pro Gly Pro Leu Gly Ser Ala
 50 55 60
 Asp Ala Gly Arg Gln Asp Leu Pro Arg Arg Val Val Arg His Glu Pro
 65 70 75 80
 Leu Gly Arg Ser Phe Leu Thr Gly Gly Leu Val Leu Leu Ala Pro Pro
 85 90 95
 Val Arg Gly Phe Gly Ala Pro Asn Ala Thr Tyr Ala Ala Arg Val Thr
 100 105 110
 Tyr Tyr Arg Leu Thr Arg Ala Cys Arg Gln Pro Ile Leu Leu Arg Gln
 115 120 125
 Tyr Gly Gly Cys Arg Gly Gly Glu Pro Pro Ser Pro Lys Thr Cys Gly
 130 135 140
 Ser Tyr Thr Tyr Thr Tyr Gln Gly Gly Gly Pro Pro Thr Arg Tyr Ala
 145 150 155 160
 Leu Val Asn Ala Ser Leu Leu Val Pro Ile Trp Asp Arg Ala Ala Glu
 165 170 175

Thr Phe Glu Tyr Gln Ile Glu Leu Gly Gly Glu Leu His Val Gly Leu
 180 185 190
 Leu Trp Val Glu Val Gly Gly Glu Gly Pro Gly Pro Thr Ala Pro Pro
 195 200 205
 Gln Ala Ala Arg Ala Glu Gly Gly Pro Cys Val Pro Pro Val Pro Ala
 210 215 220
 Gly Arg Pro Trp Arg Ser Val Pro Pro Val Trp Tyr Ser Ala Pro Asn
 225 230 235 240
 Pro Gly Phe Arg Gly Leu Arg Phe Arg Glu Arg Cys Leu Pro Pro Gln
 245 250 255
 Thr Pro Ala Ala Pro Ser Asp Leu Pro Arg Val Ala Phe Ala Pro Gln
 260 265 270
 Ser Leu Leu Val Gly Ile Thr Gly Arg Thr Phe Ile Arg Met Ala Arg
 275 280 285
 Pro Thr Glu Asp Val Gly Val Leu Pro Pro His Trp Ala Pro Gly Ala
 290 295 300
 Leu Asp Asp Gly Pro Tyr Ala Pro Phe Pro Pro Arg Pro Arg Phe Arg
 305 310 315 320
 Arg Ala Leu Arg Thr Asp Pro Glu Gly Val Asp Pro Asp Val Arg Ala
 325 330 335
 Pro Arg Thr Gly Arg Arg Leu Met Ala Leu Thr Glu Asp Thr Ser Ser
 340 345 350
 Asp Ser Pro Thr Ser Ala Pro Glu Lys Thr Pro Leu Pro Val Ser Ala
 355 360 365
 Thr Ala Met Ala Pro Ser Val Asp Pro Ser Ala Glu Pro Thr Ala Pro
 370 375 380
 Ala Thr Thr Thr Pro Pro Asp Glu Met Ala Thr Gln Ala Ala Thr Val
 385 390 395 400
 Ala Val Thr Pro Glu Glu Thr Ala Val Ala Ser Pro Pro Ala Thr Ala
 405 410 415
 Ser Val Glu Ser Ser Pro Leu Pro Ala Ala Ala Ala Thr Pro Gly
 420 425 430
 Ala Gly His Thr Asn Thr Ser Ser Ala Ser Ala Ala Lys Thr Pro Pro
 435 440 445
 Thr Thr Pro Ala Pro Thr Thr Pro Pro Pro Thr Ser Thr His Ala Thr
 450 455 460
 Pro Arg Pro Thr Thr Pro Gly Pro Gln Thr Thr Pro Pro Gly Pro Ala
 465 470 475 480

Thr Pro Gly Pro Val Gly Ala Ser Ala Ala Pro Thr Ala Asp Ser Pro
 485 490 495
 Leu Thr Ala Ser Pro Pro Ala Thr Ala Pro Gly Pro Ser Ala Ala Asn
 500 505 510
 Val Ser Val Ala Ala Thr Thr Ala Thr Pro Gly Thr Arg Gly Thr Ala
 515 520 525
 Arg Thr Pro Pro Thr Asp Pro Lys Thr His Pro His Gly Pro Ala Asp
 530 535 540
 Ala Pro Pro Gly Ser Pro Ala Pro Pro Pro Pro Glu His Arg Gly Gly
 545 550 555 560
 Pro Glu Glu Phe Glu Gly Ala Gly Asp Gly Glu Pro Pro Glu Asp Asp
 565 570 575
 Asp Ser Ala Thr Gly Leu Ala Phe Arg Thr Pro Asn Pro Asn Lys Pro
 580 585 590
 Pro Pro Ala Arg Pro Gly Pro Ile Arg Pro Thr Leu Pro Pro Gly Ile
 595 600 605
 Leu Gly Pro Leu Ala Pro Asn Thr Pro Arg Pro Pro Ala Gln Ala Pro
 610 615 620
 Ala Lys Asp Met Pro Ser Gly Pro Thr Pro Gln His Ile Pro Leu Phe
 625 630 635 640
 Trp Phe Leu Thr Ala Ser Pro Ala Leu Asp Ile Leu Phe Ile Ile Ser
 645 650 655
 Thr Thr Ile His Thr Ala Ala Phe Val Cys Leu Val Ala Leu Ala Ala
 660 665 670
 Gln Leu Trp Arg Gly Arg Ala Gly Arg Arg Arg Tyr Ala His Pro Ser
 675 680 685
 Val Arg Tyr Val Cys Leu Pro Pro Glu Arg Asp
 690 695

<210> 144
 <211> 1599
 <212> DNA
 <213> HSV2

<400> 144
 atggagctca gctatgccac caccctgcac caccgggaac ttgtgtttta cgtcacggca 60
 gacagaaacc gcgcctactt tgtgtgcggg gggtccgttt attccgtagg gcggcctcgg 120
 gattctcagc cgggggaaat tgccaagttt ggcttggtgg tccgggggac aggcccaaaa 180
 gaccgcatgg tcgccaacta cgtacgaagc gagctccgcc agcgcggcct gcgggacgtg 240
 cggcccggtg gggaggacga ggtgttcctg gacagcgtgt gtctgctaaa cccgaacgtg 300
 agctccgagc gagacgtgat taataccaac gacgttgaag tgctggacga atgcctggcc 360
 gaatactgca cctcgctgcg aaccagcccg ggggtgctgg tgaccggggt gcgctgcgc 420

```

gcgcgagaca ggggtcatcga gctattttgag caccgcggcga tcgtcaacat ttctctcgcg 480
ttcgcgtaca cccctctccc ctacgtattc gccctggccc aggcgcacct ccccggtc 540
ccgagctcgc tggagccctt ggtgagcggc ctgtttgacg gcattcccgc ccgcgcag 600
cccctggacg cccgcgaccg gcgcacggat gtcgtgatca cgggcacccg cgccccaga 660
ccgatggccg ggaccggggc cggggggcgcg ggggccaagc gggccaccgt cagcgagttc 720
gtgcaagtga agcacatcga ccgtgttggtg tcccagagcg tctcttccgc ccccccgcg 780
agcgcccccg acgcgagttt gccgcccccg gggctccagg aggcgcacct gccggggccc 840
ccgctcaggg agctgtggtg ggtgtttctac gccggcgacc gggcgctgga ggagccccac 900
gccgagtcgg gattgacgcg cgaggagggtc cgcgcctgac atgggttccg ggagcaggcg 960
tggagctgtt ttgggtcggg gggggctccg cggggcgttc tccggggccg gctggccctg 1020
agcccgaccc aaaagctcgc cgtctactac tatctcatcc accgggagcg gcgcatgtcc 1080
cccttccccg cgtcgtgctg gctcgtcggg cggtagatcc agcggcacgg cctgtacgtt 1140
cccgcgcccc acgaaccgac gttggccgat gccatgaac ggctgttccg cgacgcgctg 1200
gcggccggga ccgtggccga gcagctcctc atgttcgacc tctcccgc caaggacgtg 1260
ccggtgggga gcgacgcgcg ggccgacagc gccgcccctg tgcgtttgt ggactcgcaa 1320
cgcttgaccc cggggggggt cgtctcgccc gagcacgtca tgtacctcg cgcgcttctg 1380
ggcgtgttgt acgcgggcca cggacgcctg gcccgggcca cgcataccgc gcgcctgacg 1440
ggcgtgacgt ccctggtcct gaccgtgggg gacgtcgacc ggtgtccgc gtttgaccgc 1500
gggcggcggg gggcggtctg ccgcacgcga accgcgggtt acctggacgc gctgcttacc 1560
gtttgcctgg ctgcgcacca gcacggccag tctgtgtga 1599

```

<210> 145
 <211> 1110
 <212> DNA
 <213> HSV2

```

<400> 145
atgagtcagt ggggggcccag ggcgatcctt gtccagacgg acagcaccaa ccggaatgcc 60
gatggggact ggcaagcggc cgtagctatt cgcggggggc gagtcgttca actgaacatg 120
gtcaacaaac gcgcctgga ttttaccctg gcagaatgcg gggactccga atgggcccgtg 180
ggccgcgtct ctctgggcct gcgaatggca atgcgcggg acttctgcgc gattattcac 240
gcccccgcg tatccggccc cgggccccac gtgatgctcg gtctcgtcga ctcgggctac 300
cgcggaaccg tcctggccgt ggtcgtagcc ccgaacggga cgcgcgggtt tgcccccggg 360
gccctccggg tcgacgtgac gtttctggac atccgggcca ccccccgac cctcaccgag 420
ccgagctccc tgcaccggtt tccgcagttg gcgcgctccc cgtggcagg gttacgagaa 480
gatccttggt tggacggggc gctcgcgacc gccggggggg cggtgccctt gccggccaga 540
cggcgcgggg gatcgtggt ctacgcgggc gagctaaccg aggtgaccac cgagcacggc 600
gactgcgtgc acgagggccc cgcctttctg ccaaagcgcg aggaggacg aggccttgac 660
attctcatcc accagaccgt gaccgtcccg gccaacggcg ccacggtcat acagccgtcc 720
ctccgcgtat tgcgcgcggc cgacggacca gaggcctgct atgtgctggg gcggtcgtcg 780
ctcaatgcca ggggcctcct ggtcatgcct acgcgtggc cctccgggca cgctgtgcg 840
ttgtttgtat gtaacctgac cggagtcccg gtgacctac aagccgggtc caaggtcgcc 900
cagctgctcg tcgcggggac ccacgccctc cctggatcc ccccgacaa catccacgag 960
gacggcgcat tccgggccta cccagaggg gttccggacg cgaccgccac ccccgagac 1020
ccgccgattt tgggtgtttac gaacgagttt gacgcggacg ccccccaag caagcggggg 1080
gccggggggt ttggctccac tggcatctag 1110

```

<210> 146
 <211> 1446
 <212> DNA
 <213> HSV2

```

<400> 146
atggcctgtc gtaagttctg tggggtctac cgtagaccgc acaagagaca ggaggcgtcc 60
gtcccgccgg agacaaacac ggccccggcc tccccggcga gcacctttta taccgccgcg 120
gaggatgcgt acctggcccc cgggcccccg gaaaccatcc acccttcccg cccaccgtcc 180
cccggcgagg ctgcgcgcct gtgtcagctg caggagatct tggcccagat gcacagcgac 240
gaggactacc ccatcgtgga cgccgcgggt gcggaggagg aagacgaggc cgacgatgac 300

```



```

gccccggatg acgtggccta cccggaggac tacgcggagg ggcgtttttct gtccatgggt 360
tcggccgccc ccctgcccgg agccagcggc catcctcctg ttccggggccg cgcagccccc 420
cccgacgtcc ggacctgcga cacgggtaag gtggggggcca cggggttcac cccggaagag 480
ctcgacacca tggaccggga ggcaacttcg gccatcagcc gcgggtgcaa gcccccttcg 540
accctggcaa aactgggtgac cgggctggga ttgcgcatcc acggagcgt catcccgggg 600
tcggaggggt gtgtctttga tagcagccac ccgaactacc ctcatcgggt aatcgtcaag 660
gcgggggtgt acgccagcac gagccacgag gcgcgggtgc tgagacgcct gaaccacccc 720
gcgatcctac ccctcctgga cctgcaagtc gtttctgggg tcacgtgtct ggtcctcccc 780
aagtatcact gcgacctgta tacctatctg agcaagcgcc cgtctccgtt gggccacctt 840
cagataaccg cggctctccc gcagctcttg agcgccatcg actacgtcca ctgcaaaggc 900
atcatccacc gcgatattaa gaccgagaac atcttcatca acacccccga gaacatctgt 960
ctgggggact ttggggcggc gtgctttgtg cgcgggtgtc gatcgagccc cttccattac 1020
gggatcgag gcaccatcga taaaaacgcc ccgaggtcc tggccgggga tccgtacacc 1080
caggtaatcg acatctggag cgcggcctg gtgatctttg agaccgccgt ccacacgcg 1140
tccttgttct cggccccgcg cgaccccgaa aggcggccgt gcgacaacca gatcgcgcg 1200
atcatccgac aggccaggt acacgtcgac gagtttccga cgcacgcgga atcgcgctc 1260
accgcgcact accgctcgcg ggcgccggg aacaatcgtc cggcgtggac ccgaccggcg 1320
tggaaccgct actacaagat ccacacagac gtcgaatat tcatatgcaa agcccttacc 1380
tttgacgcgg cgctccgccc aagcgccgcg gagttgctgc gcctgccgt atttcacct 1440
aagtga 1446

```

<210> 147
<211> 1539
<212> DNA
<213> HSV2

```

<400> 147
atggtaccg acattgatat gctaatacga ctaggattgg acctgtccga cagcgagctc 60
gaggaggacg ctctggagcg ggacgaggag ggccgcccgc acgacccccg gtccgacagc 120
agcggggagt gttcctcgtc ggacgaggac atggaagacc cctgcggaga cggagggggc 180
gaggccatcg acgcggcgat tcccaaaggt cccccggccc gccccgagga cgcgggcacc 240
ccgaagcct cgacgcctcg cccggcagcg cggcggggag ccgacgatcc gccacccgcg 300
accaccggcg tgtggtcgcg cctcgggacc aggcggtcgg cttccccccg ggaaccgcac 360
ggggggaag tggcccgcat ccaacccccg tcgaccaagg caccgcatcc ccgagggcgg 420
cggcgaggtc gccgcgggg ccggggtcga tacggccccg gcggcgccga ctccacacca 480
aaaccccgcc agcgcgcttc cagaaacgcc cacaaccaag ggggtcgcca ccccgcgctc 540
gcgcggacgg acggcccccg cgccacccac ggcgaggcgc ggcgcgagg ggagcagctc 600
gacgtctccg ggggcccccg gccacgaggg acgcgccagg cccccctcc gctgatggcg 660
ctgtccctga ccccccgca cgcggacggc cgcgcccccg tcccggagcg aaaggcgccc 720
totgccgaca ccatcgaccc cgccgttcgg gcggttctgc gatccatata cgagcgcgcg 780
gcggtcgagc gcatcagcga aagctttgga cgcagtgcc ttggtcatgca agaccccttt 840
ggcgggatgc cgtttccgc cgcgacagc ccctgggctc ccgtgctggc caccgaagcg 900
ggggggtttg acgccgagac ccgtcggtt tcctgggaaa ccctggtcgc tcacggccc 960
agcctctacc gcacattcgc agccaacccg cgggcgcgt cgacagccaa ggccatgcgc 1020
gactgcgtgc tgcgccagga aaatctcacc gaggccttg cgtccgcgga tgagacgtg 1080
gcgtggtgca agatgtgcat tcaccacaat ctgcgctcc gccccagga ccctatcacc 1140
ggaacggcgg ccgcgtgct ggaaaacctc gccacgcgc tcgcgccctt tctgcagtgc 1200
tacctgaagg cccgagccct gtgcgggtg gcagacctgt gctcgcgcg acgctgtgc 1260
gacattaagg atattgcctc ctttgtgtg gtcacctgg cccgcctcgc caaccgcgtc 1320
gagcgcgcg tgtcggagat cgactacacg accgtggggg ttggggccgg cgagacgatg 1380
cacttttaca tcccgggggc ctgcatggcg ggtctcattg aaatactgga cagcaccgc 1440
caggagtgtt ccagtcgct gtgcgagctg acggccagtc acactatcgc ccccttatat 1500
gtgcacggca aatacttcta ctgcaactcc ctattttag 1539

```

<210> 148
<211> 1638
<212> DNA
<213> HSV2

<400> 148

```

atggctcgcg gggccggggt ggtgtttttt gttggagttt gggtcgtatc gtgcctggcg 60
gcagcaccga gaacgtcctg gaaacgggta acctcgggcg aggacgtggt gttgcttccg 120
gcgcccgcgg aacgcaccgg ggcccacaaa ctactgtggg ccgcggaacc cctggatgcc 180
tgcggtcccc tgcgcccgtc gtgggtggcg ctgtggcccc cccgacgggt gctcgagacg 240
gtcgtggatg cggcgtgcat gcgcgccccg gaaccgctcg ccatagcata cagtcccccg 300
ttccccgcgg gcgacgaggg actgtattcg gagttggcgt ggcgcgatcg cgtagccgtg 360
gtcaacgaga gtctgtgcat ctacggggcc ctggagacgg acagcggctc gtacaccctg 420
tccgtggtcg gcctaagcga cgaggcgcgc caagtggcgt cgggtggttct ggtcgtggag 480
cccgcacctg tgccgacccc gacccccgac gactacgacg aagaagacga cgcgggcggtg 540
acgaacgcac gccggtcagc gttccccccc caaccccccc cccgtcgtcc ccccgctgcc 600
ccccgcagcg accctcgtgt tatccccgag gtgtcccacg tgcgcggggt aacggtccat 660
atggagaccc tggaggccat tctgtttgcc cccggggaga cgtttgggac gaacgtctcc 720
atccacgcca ttgccacga cgacggtccg tacgccatgg acgtcgtctg gatgcggttt 780
gacgtgccgt cctcgtgcgc cgatatgcgg atctacgaag cttgtctgta tcaccccgag 840
cttccagagt gtctatctcc ggccgacgcg ccgtgcgcgc taagttcctg ggcgtaccgc 900
ctggcggtcc gcagctacgc cggctgttcc aggactacgc cccgcgcgcg atgttttgcc 960
gaggtcgca tggaaacgggt cccggggttg gcgtggctgg cctccaccgt caatctggaa 1020
ttccagcacg cctcccccca gcacgcccgc ctctacctgt gcgtggtgta cgtggacgat 1080
catatccacg cctggggcca catgaccatc agcacgcgcg cgagtagccg gaacgcggtg 1140
gtggaacagc acctccccca gcgccagccc gagcccgctc agccaccccg cccgcacgtg 1200
agagcccccc atccgcgcgc ctccgcgcgc ggcccgtgc gcctcggggc ggtgctgggg 1260
gcggccctgt tgcgtggcgc cctcgggctg tccgcgtggg cgtgcatgac ctgctggcgc 1320
aggcgtcctt ggcggggcgg taaaagccgg gcctcggcga cgggccccac ttacattcgc 1380
gtggcggaca gcgagctgta cgcggactgg agttcggaca gcgaggggga gcgcgacggg 1440
tccctgtggc aggacctcc ggagagaccc gactctccct ccacaaatgg atccggttt 1500
gagatcttat caccaacggc tccgtctgta tacccccata gcgagggcg taaatctcgc 1560
cgcccgtca ccacctttgg ttcgggaagc ccgggcgcgc gtcactcca ggcctcctat 1620
ccgtccgtcc tctggtaa 1638

```

<210> 149

<211> 4125

<212> DNA

<213> HSV2

<400> 149

```

atggccgctc ctgccgcga cccccgggt taccggtacg ccgcggccat cctgcccacc 60
ggctccatcc tgagtacgat cgagggtggc tcccaccgca gactctttga ttttttcgcc 120
gccgtgcgtt ccgacgaaaa cagcctgtat gacgtagagt ttgacgccct gctgggggtcc 180
tactgcaaca ccctgtcgtt cgtgcgtttt ctggagctcg gcctgtccgt ggcgtgcgtg 240
tgaccaagt tcccgagct ggcttacatg aacgaagggc gtgtgcagtt cgaggtccac 300
cagccctca tcgccgcga cggccgcac ccgtcgagc agcccggtga taattacatg 360
acgaaggtca tcgaccgcg ggccctgaac gccgccttca gcctggccac cgaggccatt 420
gccctgctca cgggggaggg cctggacggg acgggcatta gcctgcatcg ccagctgcgc 480
gccatccagc agctcgcgc caacgtccag gccgtcctgg gggcgtttga gcgcggcacg 540
gccgaccaga tgctgcacgt gctgttgag aaggcgcctc ccctggccct gctgttgccc 600
atgcaacgat atctcgacaa cgggcgcctg gcgaccaggg ttgccggggc gacctggtc 660
gccgagctga acggagctt ttgcgacag agctcttcc tgggcaaggc gggccatcgc 720
cgcgaggcca tcgaggcctg gctcgtggag ctgaccacgg cgacgcagcc gtccgtggcc 780
gtgccccgcc tgacgcacgc cgacacgcgc gggcgccggc tcgacggggg gctggtcacc 840
accgccgcca tcaaacagcg cctcctgcag tccttctga aggtggagga caccgaggcc 900
gacgtgccgg tgacctacgg cgagatggtc ttgaacgggg ccaacctcgt cacggcgtg 960
gtgatgggca aggcctgcg gagcctggac gacgtgggccc gccacctgct ggatatgcag 1020
gaggagcaac tcgaggcgaa ccgggagacg ctggatgaac tcgaaagcgc ccccagaca 1080
acgcgcgtgc gcgcggtct ggtggccata ggcgacagggc tggctcttcc ggaggccctg 1140
gagagacgca tctacgcgc caccaacgtg ccctaccccc tgggtgggcgc catggacctg 1200
acgttcgtcc tgcacctggg gctgttcaac ccggccatgg agcgcttcgc cgcgcacgcc 1260

```

```

ggggacctgg tccccgcccc cgcccaaccg gagccccgcg cgttccctcc ccggcagctg 1320
tttttttggg gaaaggacca ccaggttctg cggctgtcca tggagaacgc ggctggggacc 1380
gtgtgtcatc cttecgctcat gaacatcgac gcgcccgctg ggggcgtgaa ccacgacccc 1440
gtcgaggccg cgaatccgta cggggcgtae gtcgcggccc cggccggccc ccggcgggac 1500
atgcagcagc gttttctgaa cgcctggcgg cagcgcctcg cccacggccc ggtecggtgg 1560
gtgccgagt gccagatgac cgcggagcag ttcatgcagc ccgacaacgc caacctggct 1620
ctggagctgc accccgcgtt cgaacttctt cggggcggtg ccgacgtcga gcttcccggc 1680
ggcgaagtcc ccccgccggg tccggggggc atccaggcca cctggcgcgt ggtcaacggc 1740
aacctgcccc tggcgctgtg tccggtggcg ttctgtgacg cccggggcct ggagctcggc 1800
gttgcccgcc acgcatggc gccggctacc atagccgcg tccgcggggc gttcaggagc 1860
cgcagctacc cggcgggtgt ttacctgctg caagccgcga ttcacggcaa cgagcacgtg 1920
ttctgcgccc tggcgcggtt cgtgactcag tgcataccca gctactggaa caacacgcga 1980
tgccggcggt tcgtgaacga ctactcgctg gtctcgtaca tcgtgacct cctcgggggg 2040
gacctccccg aggagtgcac ggccgtgtat cgggacctgg tggccacgt cgaggccctg 2100
gcccagctgg tggacgactt taccctggcg ggcgcggagc tgggcgggca ggctcaggcc 2160
gagctgaatc acctgatgcg cgaccggcg ctgctgcgc cctcgtgtg ggactgcgac 2220
ggccttatgc gacacgcggc cctggaccgc caccgagact gccggattga cgcggggggg 2280
cacgagcccg tctacgcggc ggctgtcaac gtggcgacgg ccgactttaa ccgcaacgac 2340
ggccggctgc tgcacaacac ccaggcccg cgggcccagc ccgcccagca ccggccgcac 2400
cgccggcccg actggaccgt ccaccacaaa atctactatt acgtgctggt gccggccttc 2460
tcgcgggggg gctgctgcac cgcgggggtc cgttctgacc gcgtgtacgc cacgctgcag 2520
aacatggtgg tcccggagat cgcccccggt gaggagtgc cgagcgatcc cgtgaccgac 2580
cccgccacc cgctgcatcc cgccaatctg gtggccaaca cggccaagcg catgttccac 2640
aacgggcggc tcgtcgtcga cgggcccgc atgctcacgc tgcaggtgct ggccacacac 2700
atggccgagc gcacgacggc gctgctgtgc tccgcggcgc ccgacgcgg cgccaacacc 2760
gcgtcgacgg ccaacatgcg catcttcgac ggggcgctgc acgcccggct gctgctcatg 2820
gccccccagc cactggacca caccatccaa aatggcgaat acttctacgt cctgcccgtc 2880
cacgcctgt ttgcggggc cgaccacgtg gccaacgcgc ccaacttccc cccggccctg 2940
cgcgacctgg cgcgcgacgt cccctggtc ccccgggccc tggggggcaa ctacttctcc 3000
tccatccgcc agcccggtgt gcagcacgcc cgcgagagcg cggcggggga gaacgcgctg 3060
acctacgcgc tcatggcggt gtacttcaag atgagcccc tggccctgta tcaccagctc 3120
aagacggggc tccaccccg gtccgggttc accgtcgtgc ggcaggaccg ctctgtgacc 3180
gagaacgtgc tgttttccga gcgcgcgtcg gaggcgctact ttctgggcca gctccaggtg 3240
gcccgccacg aaacggggcg ggggtcaac ttacgctca cccagccgcg cggaacgtg 3300
gacctgggtg tgggctacac cgccgtcgcg gccacgggca ccgtccgcaa ccccgttacg 3360
gacatgggca acctccccca aaacttttac ctggccgcg gggccccccc gctgctagac 3420
aacgcggccg ccgtgtacct gcgcaacgcg gtctggcggt gaaaccggct gggccggcc 3480
cagcccccc cggtcttttg ctgcgccag gtgcgcgggc gcgccggcat ggaccacggg 3540
caggatgccg tgtgtgagtt catcgccacc cccgtggcca cggacatcaa ctactttcgc 3600
cgccctgca acccgcgggg acgcgcggcc ggcgcggtgt acgcggggga caaggagggg 3660
gacgtcatag cctcatgta cgaccacggc cagagcgacc cggcgcggcc ctctcgggcc 3720
acggccaacc cgtggcgctc gcagcggttc tcgtacgggg acctgctgta caacggggcc 3780
tatcacctca acggggcctc gccgtcctc agccctgct tcaagttctt caccgcggcc 3840
gacatcacgg ccaaacatcg ctgcctggag cgtctcatcg tggaaacggg atcggcggta 3900
tccacggcca ccgtgccag cgacgtgcag tttaagcgcc cgcgggggtg ccgcgagctc 3960
gtggaagacc cgtgcggcct gtttcaggaa gcctaccoga tcacctgcgc cagcgacccc 4020
gccctgctac gcagcgcccg cgatggggag gccacgcgc gagagacca ctttacgcag 4080
tatctcatct acgacgcctc cccgctaaag ggctgtctc tgtaa 4125

```

<210> 150

<211> 2169

<212> DNA

<213> HSV2

<400> 150

```

atgcaacgcc gggcgcgcg cgcgagctcc ctgcggctgg cgcgggtgcct gacgcgggcc 60
aacctcatcc gggcgcccaa cgcggcgctc ccgagcggc gcactcttcgc cgggtgtctg 120
ctccccacc cggagggggt cctcagcgcg gccgtgggg tctgcggca gcgcggccag 180

```

gacctgcagc	cggcgtttct	gaccggcgcc	gacgcgagcg	tccggctggc	ggcgcgccac	240
cataacaccg	tccccgagag	cctgatcgta	gacgggctcg	ccagcgaccc	gcactacgac	300
tacatccggc	actacgcgtc	ggccgccaaag	caggcgctcg	gcgaggtgga	gctgtcgggc	360
ggccagctga	gccgcgccat	cctagcgag	tactggaagt	acctccagac	ggctgtgccc	420
agcggcctgg	acatccccga	cgaccggcg	ggcgactgcg	accccagcct	gcacgtgctg	480
ctgcggccca	ccctgctccc	gaagctgctg	gtgcgcgccc	cgttcaagag	cggggccggc	540
gcggccaagt	acgccgccgc	ggtggcgggg	ttgcgcgacg	cggcccacag	gctccagcag	600
tacatgttct	ttatgcgccc	cgcagacccg	agccggccga	gcacggacac	cgcactgcgg	660
ctgagcgagc	tcctggccta	cgtctccgtg	ttgtaccatt	gggcctcgtg	gatgctgtgg	720
acggcggaca	agtacgtgtg	tgcgcgcctg	ggccccgcgg	atcgccggtt	cgtggcgctc	780
agcgggagtc	tggaggcgcc	cgcggagacg	tttgcgcgcc	acctggaccg	cggggcccagc	840
ggcaccacgg	gctcgatgca	gtgcatggcc	ctgcgggcg	cggtcagcga	cgtcctgggc	900
cacctgacgc	gcctggccca	cctgtgggag	accggcaagc	gcagcgcgcg	cacgtacggg	960
atcgtggagc	ccatcgcttc	gaccgtcgag	gttctatcca	tagtccacca	ccacgcccag	1020
tatataatta	acgcgacgtt	taccgggtat	gtcgtctggg	cctccgacag	cctgaacaac	1080
gagtacctta	cggcgcggtg	ggacagccag	gagcgcttct	gcaggaccgc	cgcggccctg	1140
ttccccacga	tgaccgcccc	gagctggggc	cggatggaac	tcagcatcaa	gtcctgggtt	1200
ggggcgcccc	tggccccgga	cctgcttcgg	agcggaaacc	cgtcgcccca	ctacgagtcc	1260
atcctgcgcc	tcgcggcgctc	cggccccacc	gggggcccgc	gcgcggtcgg	cgggagctgc	1320
cgggacaaga	tacaacggac	ccggcgcgac	aacgcacccc	cgcgcgtccc	ccgggctcgc	1380
ccccactcga	ccccgcggc	ccctcgagag	tgcaggcgcc	accgcgagga	cctccccgag	1440
ccccgcacg	tcgacgcggc	cgaccggggt	cccagaccct	gcgcggcgcc	gccggccacg	1500
tattacacgc	atatggccgg	ggcgcccccg	cgcctcccgc	cccgaacccc	cgcgcccccc	1560
gagcagcggc	cggcagccgc	ggcgcgcccc	ctcgcggttc	agcgcgaggc	cgcgggggtc	1620
tacgacgcgg	tgcggacctg	ggggccggac	gcggaggccg	aaccggacca	gatggaaaac	1680
acgtatctgc	tgcccagca	tgacgcgcgc	atgccgcggg	gcgtcggggt	tggcgccacc	1740
cccgcccgcc	acaccaccgc	cgcgcggccc	tggccggccg	aaagccacgc	cccccgcgcc	1800
ccctccgagg	acgcagattc	catttacgag	tgggtggggc	aggatggggg	gcgcgtctac	1860
gaggagatcc	cctgggttcg	ggtatacgaa	aacatctgcc	ctcgccggcg	tcttgccggc	1920
ggggccgccc	tgccgggaga	cgccccggac	tccccgtaca	tcgaggcgga	aaatcccctg	1980
tacgactggg	gcgggtctgc	cctcttctcc	cctcggcggg	ccacacgcgc	cccggaaccg	2040
ggactaagcc	tgtcgcccat	gcccggccgc	ccccggacca	acgcgctggc	caacgacggc	2100
ccgacgaacg	tcgcgcgcct	cagcgccctg	ttgacgaagc	tcaaacgcgg	ccgacaccag	2160
agccattaa						2169

<210> 151

<211> 957

<212> DNA

<213> HSV2

<400> 151

atgatcacgg	attgtttcga	agcagacatc	gcgatcccct	cgggtatctc	gcgccccgat	60
gccgcggcgc	tgcagcggtg	cgagggtcga	gtggtctttc	tgccgaccat	ccgcgcgcag	120
ctggcgctcg	cggacgtggc	gcacgaatcg	ttcgtctccg	gaggagttag	tcccgcacag	180
ttggggttgt	tgetggcgta	ccgcaggcgc	ttccccgcgg	taatcacgcg	ggtgctgccc	240
acgcgaatcg	tcgcctgccc	cgtggacctg	gggtcacgc	acgcgcggc	cgtcaatctc	300
cgcaacacct	ccccgctcga	cctctgcaac	ggggatccc	tcagcctcgt	cccggccgctc	360
ttcgagggcc	aggcgacgga	cgtgcgcctg	gagtcgctgg	acctcacgct	gcgggtttccg	420
gtcccgcctcc	caacgcgccct	ggcccgcgag	atagtcgcgc	ggctgggtcg	ccgggggcac	480
cgggacacct	accccgaacc	ccggacgccc	ggggagctcc	ccgacctcaa	cgtgctgtat	540
tacaacgggg	cccgctcttc	gctcgtggcc	gacgtccagc	aactcgccct	cgtaaacacc	600
gagctgcggg	cgtcgtctcc	caacatgggc	tactccataa	ccgaagggaac	caccctcatc	660
ctcacgctca	tccccggtt	gctcgcgctg	agcgcggcag	acggatacgt	gaacgcgctc	720
ctgcagatgc	agagcgtcac	gcgagaagcc	gcccagctca	tccaccccca	agcccccatg	780
ctgatgcagg	acggcgaaac	caggctgccc	ctttacgagg	cgtggtcgcc	ctgggtggcg	840
cacgcggggc	aactcgggga	catcctggcc	ctggcccccg	cggttcgggt	gtgtacgttc	900
gacggcgccc	ccgttgtgca	atccggcgac	atggcccccg	ttatccgcta	cccctga	957

<210> 152
 <211> 3066
 <212> DNA
 <213> HSV2

<400> 152

```

atggaacccc ggcccggcac gagctcccgg gcggaacccc gccccgagcg gccgcccggg 60
cagacccccg gcacgggtgag agggcgaccc ccgggtctca gggccccctt tttccccgga 120
ccacccggct gcgggttggg ggtggtcgcg ggcggtgggc tcggggggcg ggacgcttga 180
cgggggccgac ccccgggccc cttaagcggg cgggggagcc ccgtggggcg tgcgccgcc 240
cccgaccctc tggggggggcg agggagggcag ggaggagccc gagagcgggg gacagggggg 300
gagacgaggg gtcggaatcc aaaggacgca gaccaccttt ggttacggac ccctttctcc 360
cccccttccg aacaaaaagc agcggggggg gggccggggg gagggaggga cacgggggac 420
acggcgcggg ggtcccgcct cagccccgcg gccctctaaa tccccccgtg tgctttgtca 480
agcagcccg cgcggcgac gcctggggga tgtcaacga catgcagtgg ctgccagca 540
gcgactcggg ggaggagacc gaggtgggaa cgaccttcac cgcgactcca 600
cctccgaggg gggcgacacg gacacggaga tgttcgaggc gggcctgatg gacgcgcca 660
cgcccccgcc ccggcccccg gccgagcgcc agggcagccc cagccccgcg gacgcgagg 720
gatcctgtgg ggggtggccc gtgggtgagg aggaagcggg agcgggaggg gggggcgacg 780
tgtgtgccgt gtgcacggac gagatcgccc cgccctcgcg ctgccagagt tttccctgcc 840
tgcacccctt ctgcatcccg tgcataaaga cctggattcc gttgcgcaac acgtgtcccc 900
tgtgcaacac cccgggtggcg tacctgatag tgggcgtgac cgccagcggg tcgttcagca 960
ccatcccgat agtgaacgac ccccgacccc gcgtggaggc cgaggcggcc gtgcggggcg 1020
gcacggccgt ggactttatc tggacgggca acccgcgagc gggcccgcg cccctgtcgc 1080
tggggggaca caggtccgc gccctgtcgc ccaccccccc gtggcccggc acggacgacg 1140
aggacgatga cctggccgac ggtgaggcg ggccgggggtc gggcgggggg cggcgggggg 1200
tcgggcgggg gtcgggcggg ggtcgggcgg gggtcgggcg ggggtcgggc ggggtcggg 1260
cgggggtcgg gcgggggtcg ggcgggggtc gggcgggggg cgggcactaa ccggggggtc 1320
ccgtctctgt ctccctctgc agtggactac gtcccggccc ccccccgaag agcggcccg 1380
cgcgggggcg gcggtgcggg ggcgacccgc ggaacctccc agcccgccgc gacccgaccg 1440
gcgccccctg gcgcccccg gagcagcagc agcgcgggcg ccccggttgc ggcgggggtg 1500
ggatctgggt ctggggcgcg ccctgccgtc gcggccgtcg tgcgagagt ggcctctctt 1560
ccccctgcgg ccggcggggg gcgcgcgcag gcgcggcggg tgggcgaaga cgccgcggcg 1620
gcggagggca ggacgcccc cgcgagacag ccccgcgcgg ccaggagcc ccccatagtc 1680
atcagcgact ctccccgcc gtctccgcgc cgcccccggg gccccggggc gctctccttt 1740
gtctcctcct cctccgcaca ggtgtcctcg gggcccgggg ggggaggtct gccacagtgc 1800
tcggggcgcg ccgcgcgcc ccgcgcggcc gtcgccccgc gcgtccggag tccgccccgc 1860
gccgcggccg ccccggtagg gtctgcgagc ccgggcccgc gccgcccgc 1920
caagcacaga gtctgggcgg ggcaggcgcg accgacgcgc gcgggtcggg agggccgggc 2040
gcggagggag gaccgggggt ccccgcgggc accaacaccc ccggtgccgc cccccacgcc 2100
gcggaggggg cggcggcccc ccccggaag aggcgcgggt cggactcggg ccccgcggcc 2160
tcgtcctccg cctcttcttc cgccgccccg cgctcgcccc tcgccccca ggggtgggg 2220
gccaagaggg cggcgcgcgc ccgggccccg gactcggact cgggcgaccg cggccacggg 2280
ccgctcgccc cggcgtccgc gggcgccgcg ccccgctcgg cgtctccgtc gtcccaggcc 2340
gggtcgccg ccgcctcttc ctctccgcc tctcctcct cgcctcttc ctctccgcc 2400
tctcctcct ccgcctcttc ctctccgcc tctcctcct ccgcctcttc ctctccgcc 2460
tctcctcct cgggcgggg tgggtgggag gtcggtccg cgtccggcg tggggagaga 2520
agagaacct ccctcgccc ccgcgtgct gcgcgcggg ggcgaggaa gtgtgcaag 2580
aagacgcgcc acgcggaggg cggccccgag cccggggccc ggcacccggc gcccgccctc 2640
acgcgtacc tgcctatcgc ggggtctcg agcgtcgtg ccctggcgcc ttacgtgaac 2700
aagacggtca cgggggactg cctgccgtc ctggacatgg agacgggcca cataggggcc 2760
tacgtggtcc tcgtggacca gacggggaac gtggcgacc tgtgcgggc cgcggcccc 2820
gcgtggagcc gccgaccct gctccccgag cacgcgcgca actgcgtgag gcccccgac 2880
taccgacgc ccccgcgtc ggagtggaac agcctctgga tgaccccggt gggcaaatg 2940
ctctttgacc agggcaccct ggtgggcgcg ctggacttcc acggcctccg gtcgcgccac 3000
ccgtggtctc gggagcaggg cgcgcccgcg ccggccggcg acgccccgc gggccacggg 3060
gagtag                                     3066

```

<210> 153
 <211> 369
 <212> PRT
 <213> HSV2

<400> 153

Met	Ser	Gln	Trp	Gly	Pro	Arg	Ala	Ile	Leu	Val	Gln	Thr	Asp	Ser	Thr
				5					10					15	
Asn	Arg	Asn	Ala	Asp	Gly	Asp	Trp	Gln	Ala	Ala	Val	Ala	Ile	Arg	Gly
			20					25					30		
Gly	Gly	Val	Val	Gln	Leu	Asn	Met	Val	Asn	Lys	Arg	Ala	Val	Asp	Phe
		35					40					45			
Thr	Pro	Ala	Glu	Cys	Gly	Asp	Ser	Glu	Trp	Ala	Val	Gly	Arg	Val	Ser
	50					55					60				
Leu	Gly	Leu	Arg	Met	Ala	Met	Pro	Arg	Asp	Phe	Cys	Ala	Ile	Ile	His
65					70					75					80
Ala	Pro	Ala	Val	Ser	Gly	Pro	Gly	Pro	His	Val	Met	Leu	Gly	Leu	Val
				85					90					95	
Asp	Ser	Gly	Tyr	Arg	Gly	Thr	Val	Leu	Ala	Val	Val	Val	Ala	Pro	Asn
		100						105					110		
Gly	Thr	Arg	Gly	Phe	Ala	Pro	Gly	Ala	Leu	Arg	Val	Asp	Val	Thr	Phe
		115					120					125			
Leu	Asp	Ile	Arg	Ala	Thr	Pro	Pro	Thr	Leu	Thr	Glu	Pro	Ser	Ser	Leu
	130					135					140				
His	Arg	Phe	Pro	Gln	Leu	Ala	Pro	Ser	Pro	Leu	Ala	Gly	Leu	Arg	Glu
145					150					155					160
Asp	Pro	Trp	Leu	Asp	Gly	Ala	Leu	Ala	Thr	Ala	Gly	Gly	Ala	Val	Ala
			165						170					175	
Leu	Pro	Ala	Arg	Arg	Arg	Gly	Gly	Ser	Leu	Val	Tyr	Ala	Gly	Glu	Leu
		180						185					190		
Thr	Gln	Val	Thr	Thr	Glu	His	Gly	Asp	Cys	Val	His	Glu	Ala	Pro	Ala
		195					200					205			
Phe	Leu	Pro	Lys	Arg	Glu	Glu	Asp	Ala	Gly	Phe	Asp	Ile	Leu	Ile	His
	210					215					220				
Arg	Ala	Val	Thr	Val	Pro	Ala	Asn	Gly	Ala	Thr	Val	Ile	Gln	Pro	Ser
225					230					235					240
Leu	Arg	Val	Leu	Arg	Ala	Ala	Asp	Gly	Pro	Glu	Ala	Cys	Tyr	Val	Leu
			245						250					255	
Gly	Arg	Ser	Ser	Leu	Asn	Ala	Arg	Gly	Leu	Leu	Val	Met	Pro	Thr	Arg
		260						265					270		

Trp Pro Ser Gly His Ala Cys Ala Phe Val Val Cys Asn Leu Thr Gly
 275 280 285
 Val Pro Val Thr Leu Gln Ala Gly Ser Lys Val Ala Gln Leu Leu Val
 290 295 300
 Ala Gly Thr His Ala Leu Pro Trp Ile Pro Pro Asp Asn Ile His Glu
 305 310 315 320
 Asp Gly Ala Phe Arg Ala Tyr Pro Arg Gly Val Pro Asp Ala Thr Ala
 325 330 335
 Thr Pro Arg Asp Pro Pro Ile Leu Val Phe Thr Asn Glu Phe Asp Ala
 340 345 350
 Asp Ala Pro Pro Ser Lys Arg Gly Ala Gly Gly Phe Gly Ser Thr Gly
 355 360 365

Ile

<210> 154
 <211> 532
 <212> PRT
 <213> HSV2

<400> 154

Met Glu Leu Ser Tyr Ala Thr Thr Leu His His Arg Asp Val Val Phe
 5 10 15
 Tyr Val Thr Ala Asp Arg Asn Arg Ala Tyr Phe Val Cys Gly Gly Ser
 20 25 30
 Val Tyr Ser Val Gly Arg Pro Arg Asp Ser Gln Pro Gly Glu Ile Ala
 35 40 45
 Lys Phe Gly Leu Val Val Arg Gly Thr Gly Pro Lys Asp Arg Met Val
 50 55 60
 Ala Asn Tyr Val Arg Ser Glu Leu Arg Gln Arg Gly Leu Arg Asp Val
 65 70 75 80
 Arg Pro Val Gly Glu Asp Glu Val Phe Leu Asp Ser Val Cys Leu Leu
 85 90 95
 Asn Pro Asn Val Ser Ser Glu Arg Asp Val Ile Asn Thr Asn Asp Val
 100 105 110
 Glu Val Leu Asp Glu Cys Leu Ala Glu Tyr Cys Thr Ser Leu Arg Thr
 115 120 125
 Ser Pro Gly Val Leu Val Thr Gly Val Arg Val Arg Ala Arg Asp Arg
 130 135 140
 Val Ile Glu Leu Phe Glu His Pro Ala Ile Val Asn Ile Ser Ser Arg
 145 150 155 160

Phe Ala Tyr Thr Pro Ser Pro Tyr Val Phe Ala Leu Ala Gln Ala His
 165 170 175
 Leu Pro Arg Leu Pro Ser Ser Leu Glu Pro Leu Val Ser Gly Leu Phe
 180 185 190
 Asp Gly Ile Pro Ala Pro Arg Gln Pro Leu Asp Ala Arg Asp Arg Arg
 195 200 205
 Thr Asp Val Val Ile Thr Gly Thr Arg Ala Pro Arg Pro Met Ala Gly
 210 215 220
 Thr Gly Ala Gly Gly Ala Gly Ala Lys Arg Ala Thr Val Ser Glu Phe
 225 230 235 240
 Val Gln Val Lys His Ile Asp Arg Val Val Ser Pro Ser Val Ser Ser
 245 250 255
 Ala Pro Pro Pro Ser Ala Pro Asp Ala Ser Leu Pro Pro Pro Gly Leu
 260 265 270
 Gln Glu Ala Ala Pro Pro Gly Pro Pro Leu Arg Glu Leu Trp Trp Val
 275 280 285
 Phe Tyr Ala Gly Asp Arg Ala Leu Glu Glu Pro His Ala Glu Ser Gly
 290 295 300
 Leu Thr Arg Glu Glu Val Arg Ala Val His Gly Phe Arg Glu Gln Ala
 305 310 315 320
 Trp Lys Leu Phe Gly Ser Val Gly Ala Pro Arg Ala Phe Leu Gly Ala
 325 330 335
 Ala Leu Ala Leu Ser Pro Thr Gln Lys Leu Ala Val Tyr Tyr Tyr Leu
 340 345 350
 Ile His Arg Glu Arg Arg Met Ser Pro Phe Pro Ala Leu Val Arg Leu
 355 360 365
 Val Gly Arg Tyr Ile Gln Arg His Gly Leu Tyr Val Pro Ala Pro Asp
 370 375 380
 Glu Pro Thr Leu Ala Asp Ala Met Asn Gly Leu Phe Arg Asp Ala Leu
 385 390 395 400
 Ala Ala Gly Thr Val Ala Glu Gln Leu Leu Met Phe Asp Leu Leu Pro
 405 410 415
 Pro Lys Asp Val Pro Val Gly Ser Asp Ala Arg Ala Asp Ser Ala Ala
 420 425 430
 Leu Leu Arg Phe Val Asp Ser Gln Arg Leu Thr Pro Gly Gly Ser Val
 435 440 445
 Ser Pro Glu His Val Met Tyr Leu Gly Ala Phe Leu Gly Val Leu Tyr
 450 455 460
 Ala Gly His Gly Arg Leu Ala Ala Ala Thr His Thr Ala Arg Leu Thr

465 470 475 480
 Gly Val Thr Ser Leu Val Leu Thr Val Gly Asp Val Asp Arg Met Ser
 485 490 495
 Ala Phe Asp Arg Gly Pro Ala Gly Ala Ala Gly Arg Thr Arg Thr Ala
 500 505 510
 Gly Tyr Leu Asp Ala Leu Leu Thr Val Cys Leu Ala Arg Ala Gln His
 515 520 525
 Gly Gln Ser Val
 530

<210> 155
 <211> 481
 <212> PRT
 <213> HSV2

<400> 155
 Met Ala Cys Arg Lys Phe Cys Gly Val Tyr Arg Arg Pro Asp Lys Arg
 5 10 15
 Gln Glu Ala Ser Val Pro Pro Glu Thr Asn Thr Ala Pro Ala Phe Pro
 20 25 30
 Ala Ser Thr Phe Tyr Thr Pro Ala Glu Asp Ala Tyr Leu Ala Pro Gly
 35 40 45
 Pro Pro Glu Thr Ile His Pro Ser Arg Pro Pro Ser Pro Gly Glu Ala
 50 55 60
 Ala Arg Leu Cys Gln Leu Gln Glu Ile Leu Ala Gln Met His Ser Asp
 65 70 75 80
 Glu Asp Tyr Pro Ile Val Asp Ala Ala Gly Ala Glu Glu Glu Asp Glu
 85 90 95
 Ala Asp Asp Asp Ala Pro Asp Asp Val Ala Tyr Pro Glu Asp Tyr Ala
 100 105 110
 Glu Gly Arg Phe Leu Ser Met Val Ser Ala Ala Pro Leu Pro Gly Ala
 115 120 125
 Ser Gly His Pro Pro Val Pro Gly Arg Ala Ala Pro Pro Asp Val Arg
 130 135 140
 Thr Cys Asp Thr Gly Lys Val Gly Ala Thr Gly Phe Thr Pro Glu Glu
 145 150 155 160
 Leu Asp Thr Met Asp Arg Glu Ala Leu Arg Ala Ile Ser Arg Gly Cys
 165 170 175
 Lys Pro Pro Ser Thr Leu Ala Lys Leu Val Thr Gly Leu Gly Phe Ala
 180 185 190
 Ile His Gly Ala Leu Ile Pro Gly Ser Glu Gly Cys Val Phe Asp Ser

195	200	205
Ser His Pro Asn Tyr Pro His Arg Val Ile Val Lys Ala Gly Trp Tyr		
210	215	220
Ala Ser Thr Ser His Glu Ala Arg Leu Leu Arg Arg Leu Asn His Pro		
225	230	235 240
Ala Ile Leu Pro Leu Leu Asp Leu His Val Val Ser Gly Val Thr Cys		
	245	250 255
Leu Val Leu Pro Lys Tyr His Cys Asp Leu Tyr Thr Tyr Leu Ser Lys		
	260	265 270
Arg Pro Ser Pro Leu Gly His Leu Gln Ile Thr Ala Val Ser Arg Gln		
	275	280 285
Leu Leu Ser Ala Ile Asp Tyr Val His Cys Lys Gly Ile Ile His Arg		
	290	295 300
Asp Ile Lys Thr Glu Asn Ile Phe Ile Asn Thr Pro Glu Asn Ile Cys		
305	310	315 320
Leu Gly Asp Phe Gly Ala Ala Cys Phe Val Arg Gly Cys Arg Ser Ser		
	325	330 335
Pro Phe His Tyr Gly Ile Ala Gly Thr Ile Asp Thr Asn Ala Pro Glu		
	340	345 350
Val Leu Ala Gly Asp Pro Tyr Thr Gln Val Ile Asp Ile Trp Ser Ala		
	355	360 365
Gly Leu Val Ile Phe Glu Thr Ala Val His Thr Ala Ser Leu Phe Ser		
	370	375 380
Ala Pro Arg Asp Pro Glu Arg Arg Pro Cys Asp Asn Gln Ile Ala Arg		
385	390	395 400
Ile Ile Arg Gln Ala Gln Val His Val Asp Glu Phe Pro Thr His Ala		
	405	410 415
Glu Ser Arg Leu Thr Ala His Tyr Arg Ser Arg Ala Ala Gly Asn Asn		
	420	425 430
Arg Pro Ala Trp Thr Arg Pro Ala Trp Thr Arg Tyr Tyr Lys Ile His		
	435	440 445
Thr Asp Val Glu Tyr Leu Ile Cys Lys Ala Leu Thr Phe Asp Ala Ala		
	450	455 460
Leu Arg Pro Ser Ala Ala Glu Leu Leu Arg Leu Pro Leu Phe His Pro		
465	470	475 480
Lys		

Met Ala Thr Asp Ile Asp Met Leu Ile Asp Leu Gly Leu Asp Leu Ser
5 10 15

Asp Ser Glu Leu Glu Glu Asp Ala Leu Glu Arg Asp Glu Glu Gly Arg
20 25 30

Arg Asp Asp Pro Glu Ser Asp Ser Ser Gly Glu Cys Ser Ser Ser Asp
35 40 45

Glu Asp Met Glu Asp Pro Cys Gly Asp Gly Gly Ala Glu Ala Ile Asp
50 55 60

Ala Ala Ile Pro Lys Gly Pro Pro Ala Arg Pro Glu Asp Ala Gly Thr
65 70 75 80

Pro Glu Ala Ser Thr Pro Arg Pro Ala Ala Arg Arg Gly Ala Asp Asp
85 90 95

Pro Pro Pro Ala Thr Thr Gly Val Trp Ser Arg Leu Gly Thr Arg Arg
100 105 110

Ser Ala Ser Pro Arg Glu Pro His Gly Gly Lys Val Ala Arg Ile Gln
 115 120 125

Pro Pro Ser Thr Lys Ala Pro His Pro Arg Gly Gly Arg Arg Gly Arg
130 135 140

Arg Arg Gly Arg Gly Arg Tyr Gly Pro Gly Gly Ala Asp Ser Thr Pro
145 150 155 160

Lys Pro Arg Arg Arg Val Ser Arg Asn Ala His Asn Gln Gly Gly Arg
165 170 175

His Pro Ala Ser Ala Arg Thr Asp Gly Pro Gly Ala Thr His Gly Glu
180 185 190

Ala Arg Arg Gly Gly Glu Gln Leu Asp Val Ser Gly Gly Pro Arg Pro
195 200 205

Arg Gly Thr Arg Gln Ala Pro Pro Pro Leu Met Ala Leu Ser Leu Thr
210 215 220

Pro Pro His Ala Asp Gly Arg Ala Pro Val Pro Glu Arg Lys Ala Pro
225 230 235 240

Ser Ala Asp Thr Ile Asp Pro Ala Val Arg Ala Val Leu Arg Ser Ile
245 250 255

Ser Glu Arg Ala Ala Val Glu Arg Ile Ser Glu Ser Phe Gly Arg Ser
260 265 270

Ala Leu Val Met Gln Asp Pro Phe Gly Gly Met Pro Phe Pro Ala Ala
275 280 285

Asn Ser Pro Trp Ala Pro Val Leu Ala Thr Gln Ala Gly Gly Phe Asp
 290 295 300
 Ala Glu Thr Arg Arg Val Ser Trp Glu Thr Leu Val Ala His Gly Pro
 305 310 315 320
 Ser Leu Tyr Arg Thr Phe Ala Ala Asn Pro Arg Ala Ala Ser Thr Ala
 325 330 335
 Lys Ala Met Arg Asp Cys Val Leu Arg Gln Glu Asn Leu Ile Glu Ala
 340 345 350
 Leu Ala Ser Ala Asp Glu Thr Leu Ala Trp Cys Lys Met Cys Ile His
 355 360 365
 His Asn Leu Pro Leu Arg Pro Gln Asp Pro Ile Ile Gly Thr Ala Ala
 370 375 380
 Ala Val Leu Glu Asn Leu Ala Thr Arg Leu Arg Pro Phe Leu Gln Cys
 385 390 395 400
 Tyr Leu Lys Ala Arg Gly Leu Cys Gly Leu Asp Asp Leu Cys Ser Arg
 405 410 415
 Arg Arg Leu Ser Asp Ile Lys Asp Ile Ala Ser Phe Val Leu Val Ile
 420 425 430
 Leu Ala Arg Leu Ala Asn Arg Val Glu Arg Gly Val Ser Glu Ile Asp
 435 440 445
 Tyr Thr Thr Val Gly Val Gly Ala Gly Glu Thr Met His Phe Tyr Ile
 450 455 460
 Pro Gly Ala Cys Met Ala Gly Leu Ile Glu Ile Leu Asp Thr His Arg
 465 470 475 480
 Gln Glu Cys Ser Ser Arg Val Cys Glu Leu Thr Ala Ser His Thr Ile
 485 490 495
 Ala Pro Leu Tyr Val His Gly Lys Tyr Phe Tyr Cys Asn Ser Leu Phe
 500 505 510

<210> 157

<211> 545

<212> PRT

<213> HSV2

<400> 157

Met Ala Arg Gly Ala Gly Leu Val Phe Phe Val Gly Val Trp Val Val
 5 10 15

Ser Cys Leu Ala Ala Ala Pro Arg Thr Ser Trp Lys Arg Val Thr Ser
 20 25 30

Gly Glu Asp Val Val Leu Leu Pro Ala Pro Ala Glu Arg Thr Arg Ala
 35 40 45

His Lys Leu Leu Trp Ala Ala Glu Pro Leu Asp Ala Cys Gly Pro Leu
 50 55 60
 Arg Pro Ser Trp Val Ala Leu Trp Pro Pro Arg Arg Val Leu Glu Thr
 65 70 75 80
 Val Val Asp Ala Ala Cys Met Arg Ala Pro Glu Pro Leu Ala Ile Ala
 85 90 95
 Tyr Ser Pro Pro Phe Pro Ala Gly Asp Glu Gly Leu Tyr Ser Glu Leu
 100 105 110
 Ala Trp Arg Asp Arg Val Ala Val Val Asn Glu Ser Leu Val Ile Tyr
 115 120 125
 Gly Ala Leu Glu Thr Asp Ser Gly Leu Tyr Thr Leu Ser Val Val Gly
 130 135 140
 Leu Ser Asp Glu Ala Arg Gln Val Ala Ser Val Val Leu Val Val Glu
 145 150 155 160
 Pro Ala Pro Val Pro Thr Pro Thr Pro Asp Asp Tyr Asp Glu Glu Asp
 165 170 175
 Asp Ala Gly Val Thr Asn Ala Arg Arg Ser Ala Phe Pro Pro Gln Pro
 180 185 190
 Pro Pro Arg Arg Pro Pro Val Ala Pro Pro Thr His Pro Arg Val Ile
 195 200 205
 Pro Glu Val Ser His Val Arg Gly Val Thr Val His Met Glu Thr Leu
 210 215 220
 Glu Ala Ile Leu Phe Ala Pro Gly Glu Thr Phe Gly Thr Asn Val Ser
 225 230 235 240
 Ile His Ala Ile Ala His Asp Asp Gly Pro Tyr Ala Met Asp Val Val
 245 250 255
 Trp Met Arg Phe Asp Val Pro Ser Ser Cys Ala Asp Met Arg Ile Tyr
 260 265 270
 Glu Ala Cys Leu Tyr His Pro Gln Leu Pro Glu Cys Leu Ser Pro Ala
 275 280 285
 Asp Ala Pro Cys Ala Val Ser Ser Trp Ala Tyr Arg Leu Ala Val Arg
 290 295 300
 Ser Tyr Ala Gly Cys Ser Arg Thr Thr Pro Pro Pro Arg Cys Phe Ala
 305 310 315 320
 Glu Ala Arg Met Glu Pro Val Pro Gly Leu Ala Trp Leu Ala Ser Thr
 325 330 335
 Val Asn Leu Glu Phe Gln His Ala Ser Pro Gln His Ala Gly Leu Tyr
 340 345 350

Leu Cys Val Val Tyr Val Asp Asp His Ile His Ala Trp Gly His Met
 355 360 365
 Thr Ile Ser Thr Ala Ala Gln Tyr Arg Asn Ala Val Val Glu Gln His
 370 375 380
 Leu Pro Gln Arg Gln Pro Glu Pro Val Glu Pro Thr Arg Pro His Val
 385 390 395 400
 Arg Ala Pro His Pro Ala Pro Ser Ala Arg Gly Pro Leu Arg Leu Gly
 405 410 415
 Ala Val Leu Gly Ala Ala Leu Leu Leu Ala Ala Leu Gly Leu Ser Ala
 420 425 430
 Trp Ala Cys Met Thr Cys Trp Arg Arg Arg Ser Trp Arg Ala Val Lys
 435 440 445
 Ser Arg Ala Ser Ala Thr Gly Pro Thr Tyr Ile Arg Val Ala Asp Ser
 450 455 460
 Glu Leu Tyr Ala Asp Trp Ser Ser Asp Ser Glu Gly Glu Arg Asp Gly
 465 470 475 480
 Ser Leu Trp Gln Asp Pro Pro Glu Arg Pro Asp Ser Pro Ser Thr Asn
 485 490 495
 Gly Ser Gly Phe Glu Ile Leu Ser Pro Thr Ala Pro Ser Val Tyr Pro
 500 505 510
 His Ser Glu Gly Arg Lys Ser Arg Arg Pro Leu Thr Thr Phe Gly Ser
 515 520 525
 Gly Ser Pro Gly Arg Arg His Ser Gln Ala Ser Tyr Pro Ser Val Leu
 530 535 540
 Trp
 545

<210> 158
 <211> 1374
 <212> PRT
 <213> HSV2

<400> 158
 Met Ala Ala Pro Ala Arg Asp Pro Pro Gly Tyr Arg Tyr Ala Ala Ala
 5 10 15
 Ile Leu Pro Thr Gly Ser Ile Leu Ser Thr Ile Glu Val Ala Ser His
 20 25 30
 Arg Arg Leu Phe Asp Phe Phe Ala Ala Val Arg Ser Asp Glu Asn Ser
 35 40 45
 Leu Tyr Asp Val Glu Phe Asp Ala Leu Leu Gly Ser Tyr Cys Asn Thr
 50 55 60

Leu Ser Leu Val Arg Phe Leu Glu Leu Gly Leu Ser Val Ala Cys Val
 65 70 75 80
 Cys Thr Lys Phe Pro Glu Leu Ala Tyr Met Asn Glu Gly Arg Val Gln
 85 90 95
 Phe Glu Val His Gln Pro Leu Ile Ala Arg Asp Gly Pro His Pro Val
 100 105 110
 Glu Gln Pro Val His Asn Tyr Met Thr Lys Val Ile Asp Arg Arg Ala
 115 120 125
 Leu Asn Ala Ala Phe Ser Leu Ala Thr Glu Ala Ile Ala Leu Leu Thr
 130 135 140
 Gly Glu Ala Leu Asp Gly Thr Gly Ile Ser Leu His Arg Gln Leu Arg
 145 150 155 160
 Ala Ile Gln Gln Leu Ala Arg Asn Val Gln Ala Val Leu Gly Ala Phe
 165 170 175
 Glu Arg Gly Thr Ala Asp Gln Met Leu His Val Leu Leu Glu Lys Ala
 180 185 190
 Pro Pro Leu Ala Leu Leu Leu Pro Met Gln Arg Tyr Leu Asp Asn Gly
 195 200 205
 Arg Leu Ala Thr Arg Val Ala Arg Ala Thr Leu Val Ala Glu Leu Lys
 210 215 220
 Arg Ser Phe Cys Asp Thr Ser Phe Phe Leu Gly Lys Ala Gly His Arg
 225 230 235 240
 Arg Glu Ala Ile Glu Ala Trp Leu Val Asp Leu Thr Thr Ala Thr Gln
 245 250 255
 Pro Ser Val Ala Val Pro Arg Leu Thr His Ala Asp Thr Arg Gly Arg
 260 265 270
 Pro Val Asp Gly Val Leu Val Thr Thr Ala Ala Ile Lys Gln Arg Leu
 275 280 285
 Leu Gln Ser Phe Leu Lys Val Glu Asp Thr Glu Ala Asp Val Pro Val
 290 295 300
 Thr Tyr Gly Glu Met Val Leu Asn Gly Ala Asn Leu Val Thr Ala Leu
 305 310 315 320
 Val Met Gly Lys Ala Val Arg Ser Leu Asp Asp Val Gly Arg His Leu
 325 330 335
 Leu Asp Met Gln Glu Glu Gln Leu Glu Ala Asn Arg Glu Thr Leu Asp
 340 345 350
 Glu Leu Glu Ser Ala Pro Gln Thr Thr Arg Val Arg Ala Asp Leu Val
 355 360 365
 Ala Ile Gly Asp Arg Leu Val Phe Leu Glu Ala Leu Glu Arg Arg Ile

370	375	380
Tyr Ala Ala Thr Asn Val Pro Tyr Pro Leu Val Gly Ala Met Asp Leu		
385	390	395 400
Thr Phe Val Leu Pro Leu Gly Leu Phe Asn Pro Ala Met Glu Arg Phe		
	405	410 415
Ala Ala His Ala Gly Asp Leu Val Pro Ala Pro Gly His Pro Glu Pro		
	420	425 430
Arg Ala Phe Pro Pro Arg Gln Leu Phe Phe Trp Gly Lys Asp His Gln		
	435	440 445
Val Leu Arg Leu Ser Met Glu Asn Ala Val Gly Thr Val Cys His Pro		
	450	455 460
Ser Leu Met Asn Ile Asp Ala Ala Val Gly Gly Val Asn His Asp Pro		
	465	470 475 480
Val Glu Ala Ala Asn Pro Tyr Gly Ala Tyr Val Ala Ala Pro Ala Gly		
	485	490 495
Pro Gly Ala Asp Met Gln Gln Arg Phe Leu Asn Ala Trp Arg Gln Arg		
	500	505 510
Leu Ala His Gly Arg Val Arg Trp Val Ala Glu Cys Gln Met Thr Ala		
	515	520 525
Glu Gln Phe Met Gln Pro Asp Asn Ala Asn Leu Ala Leu Glu Leu His		
	530	535 540
Pro Ala Phe Asp Phe Phe Ala Gly Val Ala Asp Val Glu Leu Pro Gly		
	545	550 555 560
Gly Glu Val Pro Pro Ala Gly Pro Gly Ala Ile Gln Ala Thr Trp Arg		
	565	570 575
Val Val Asn Gly Asn Leu Pro Leu Ala Leu Cys Pro Val Ala Phe Arg		
	580	585 590
Asp Ala Arg Gly Leu Glu Leu Gly Val Gly Arg His Ala Met Ala Pro		
	595	600 605
Ala Thr Ile Ala Ala Val Arg Gly Ala Phe Glu Asp Arg Ser Tyr Pro		
	610	615 620
Ala Val Phe Tyr Leu Leu Gln Ala Ala Ile His Gly Asn Glu His Val		
	625	630 635 640
Phe Cys Ala Leu Ala Arg Leu Val Thr Gln Cys Ile Thr Ser Tyr Trp		
	645	650 655
Asn Asn Thr Arg Cys Ala Ala Phe Val Asn Asp Tyr Ser Leu Val Ser		
	660	665 670
Tyr Ile Val Thr Tyr Leu Gly Gly Asp Leu Pro Glu Glu Cys Met Ala		
	675	680 685

Val Tyr Arg Asp Leu Val Ala His Val Glu Ala Leu Ala Gln Leu Val
 690 695 700
 Asp Asp Phe Thr Leu Pro Gly Pro Glu Leu Gly Gly Gln Ala Gln Ala
 705 710 715 720
 Glu Leu Asn His Leu Met Arg Asp Pro Ala Leu Leu Pro Pro Leu Val
 725 730 735
 Trp Asp Cys Asp Gly Leu Met Arg His Ala Ala Leu Asp Arg His Arg
 740 745 750
 Asp Cys Arg Ile Asp Ala Gly Gly His Glu Pro Val Tyr Ala Ala Ala
 755 760 765
 Cys Asn Val Ala Thr Ala Asp Phe Asn Arg Asn Asp Gly Arg Leu Leu
 770 775 780
 His Asn Thr Gln Ala Arg Ala Ala Asp Ala Ala Asp Asp Arg Pro His
 785 790 795 800
 Arg Pro Ala Asp Trp Thr Val His His Lys Ile Tyr Tyr Tyr Val Leu
 805 810 815
 Val Pro Ala Phe Ser Arg Gly Arg Cys Cys Thr Ala Gly Val Arg Phe
 820 825 830
 Asp Arg Val Tyr Ala Thr Leu Gln Asn Met Val Val Pro Glu Ile Ala
 835 840 845
 Pro Gly Glu Glu Cys Pro Ser Asp Pro Val Thr Asp Pro Ala His Pro
 850 855 860
 Leu His Pro Ala Asn Leu Val Ala Asn Thr Val Lys Arg Met Phe His
 865 870 875 880
 Asn Gly Arg Val Val Val Asp Gly Pro Ala Met Leu Thr Leu Gln Val
 885 890 895
 Leu Ala His Asn Met Ala Glu Arg Thr Thr Ala Leu Leu Cys Ser Ala
 900 905 910
 Ala Pro Asp Ala Gly Ala Asn Thr Ala Ser Thr Ala Asn Met Arg Ile
 915 920 925
 Phe Asp Gly Ala Leu His Ala Gly Val Leu Leu Met Ala Pro Gln His
 930 935 940
 Leu Asp His Thr Ile Gln Asn Gly Glu Tyr Phe Tyr Val Leu Pro Val
 945 950 955 960
 His Ala Leu Phe Ala Gly Ala Asp His Val Ala Asn Ala Pro Asn Phe
 965 970 975
 Pro Pro Ala Leu Arg Asp Leu Ala Arg Asp Val Pro Leu Val Pro Pro
 980 985 990

Ala Leu Gly Ala Asn Tyr Phe Ser Ser Ile Arg Gln Pro Val Val Gln
 995 1000 1005
 His Ala Arg Glu Ser Ala Ala Gly Glu Asn Ala Leu Thr Tyr Ala Leu
 1010 1015 1020
 Met Ala Gly Tyr Phe Lys Met Ser Pro Val Ala Leu Tyr His Gln Leu
 1025 1030 1035 1040
 Lys Thr Gly Leu His Pro Gly Phe Gly Phe Thr Val Val Arg Gln Asp
 1045 1050 1055
 Arg Phe Val Thr Glu Asn Val Leu Phe Ser Glu Arg Ala Ser Glu Ala
 1060 1065 1070
 Tyr Phe Leu Gly Gln Leu Gln Val Ala Arg His Glu Thr Gly Gly Gly
 1075 1080 1085
 Val Asn Phe Thr Leu Thr Gln Pro Arg Gly Asn Val Asp Leu Gly Val
 1090 1095 1100
 Gly Tyr Thr Ala Val Ala Ala Thr Gly Thr Val Arg Asn Pro Val Thr
 1105 1110 1115 1120
 Asp Met Gly Asn Leu Pro Gln Asn Phe Tyr Leu Gly Arg Gly Ala Pro
 1125 1130 1135
 Pro Leu Leu Asp Asn Ala Ala Ala Val Tyr Leu Arg Asn Ala Val Val
 1140 1145 1150
 Ala Gly Asn Arg Leu Gly Pro Ala Gln Pro Leu Pro Val Phe Gly Cys
 1155 1160 1165
 Ala Gln Val Pro Arg Arg Ala Gly Met Asp His Gly Gln Asp Ala Val
 1170 1175 1180
 Cys Glu Phe Ile Ala Thr Pro Val Ala Thr Asp Ile Asn Tyr Phe Arg
 1185 1190 1195 1200
 Arg Pro Cys Asn Pro Arg Gly Arg Ala Ala Gly Gly Val Tyr Ala Gly
 1205 1210 1215
 Asp Lys Glu Gly Asp Val Ile Ala Leu Met Tyr Asp His Gly Gln Ser
 1220 1225 1230
 Asp Pro Ala Arg Pro Phe Ala Ala Thr Ala Asn Pro Trp Ala Ser Gln
 1235 1240 1245
 Arg Phe Ser Tyr Gly Asp Leu Leu Tyr Asn Gly Ala Tyr His Leu Asn
 1250 1255 1260
 Gly Ala Ser Pro Val Leu Ser Pro Cys Phe Lys Phe Phe Thr Ala Ala
 1265 1270 1275 1280
 Asp Ile Thr Ala Lys His Arg Cys Leu Glu Arg Leu Ile Val Glu Thr
 1285 1290 1295
 Gly Ser Ala Val Ser Thr Ala Thr Ala Ala Ser Asp Val Gln Phe Lys

1300 1305 1310
 Arg Pro Pro Gly Cys Arg Glu Leu Val Glu Asp Pro Cys Gly Leu Phe
 1315 1320 1325
 Gln Glu Ala Tyr Pro Ile Thr Cys Ala Ser Asp Pro Ala Leu Leu Arg
 1330 1335 1340
 Ser Ala Arg Asp Gly Glu Ala His Ala Arg Glu Thr His Phe Thr Gln
 1345 1350 1355 1360
 Tyr Leu Ile Tyr Asp Ala Ser Pro Leu Lys Gly Leu Ser Leu
 1365 1370

 <210> 159
 <211> 722
 <212> PRT
 <213> HSV2

 <400> 159
 Met Gln Arg Arg Ala Arg Gly Ala Ser Ser Leu Arg Leu Ala Arg Cys
 5 10 15
 Leu Thr Pro Ala Asn Leu Ile Arg Gly Ala Asn Ala Gly Val Pro Glu
 20 25 30
 Arg Arg Ile Phe Ala Gly Cys Leu Leu Pro Thr Pro Glu Gly Leu Leu
 35 40 45
 Ser Ala Ala Val Gly Val Leu Arg Gln Arg Ala Asp Asp Leu Gln Pro
 50 55 60
 Ala Phe Leu Thr Gly Ala Asp Arg Ser Val Arg Leu Ala Ala Arg His
 65 70 75 80
 His Asn Thr Val Pro Glu Ser Leu Ile Val Asp Gly Leu Ala Ser Asp
 85 90 95
 Pro His Tyr Asp Tyr Ile Arg His Tyr Ala Ser Ala Ala Lys Gln Ala
 100 105 110
 Leu Gly Glu Val Glu Leu Ser Gly Gly Gln Leu Ser Arg Ala Ile Leu
 115 120 125
 Ala Gln Tyr Trp Lys Tyr Leu Gln Thr Val Val Pro Ser Gly Leu Asp
 130 135 140
 Ile Pro Asp Asp Pro Ala Gly Asp Cys Asp Pro Ser Leu His Val Leu
 145 150 155 160
 Leu Arg Pro Thr Leu Leu Pro Lys Leu Leu Val Arg Ala Pro Phe Lys
 165 170 175
 Ser Gly Ala Ala Ala Ala Lys Tyr Ala Ala Ala Val Ala Gly Leu Arg
 180 185 190
 Asp Ala Ala His Arg Leu Gln Gln Tyr Met Phe Phe Met Arg Pro Ala

195					200					205					
Asp	Pro	Ser	Arg	Pro	Ser	Thr	Asp	Thr	Ala	Leu	Arg	Leu	Ser	Glu	Leu
210							215				220				
Leu	Ala	Tyr	Val	Ser	Val	Leu	Tyr	His	Trp	Ala	Ser	Trp	Met	Leu	Trp
225					230					235					240
Thr	Ala	Asp	Lys	Tyr	Val	Cys	Arg	Arg	Leu	Gly	Pro	Ala	Asp	Arg	Arg
				245					250					255	
Phe	Val	Ala	Leu	Ser	Gly	Ser	Leu	Glu	Ala	Pro	Ala	Glu	Thr	Phe	Ala
			260					265					270		
Arg	His	Leu	Asp	Arg	Gly	Pro	Ser	Gly	Thr	Thr	Gly	Ser	Met	Gln	Cys
		275					280					285			
Met	Ala	Leu	Arg	Ala	Ala	Val	Ser	Asp	Val	Leu	Gly	His	Leu	Thr	Arg
290					295					300					
Leu	Ala	His	Leu	Trp	Glu	Thr	Gly	Lys	Arg	Ser	Gly	Gly	Thr	Tyr	Gly
305					310					315					320
Ile	Val	Asp	Ala	Ile	Val	Ser	Thr	Val	Glu	Val	Leu	Ser	Ile	Val	His
				325					330					335	
His	His	Ala	Gln	Tyr	Ile	Ile	Asn	Ala	Thr	Leu	Thr	Gly	Tyr	Val	Val
			340				345						350		
Trp	Ala	Ser	Asp	Ser	Leu	Asn	Asn	Glu	Tyr	Leu	Thr	Ala	Ala	Val	Asp
		355					360					365			
Ser	Gln	Glu	Arg	Phe	Cys	Arg	Thr	Ala	Ala	Pro	Leu	Phe	Pro	Thr	Met
370					375					380					
Thr	Ala	Pro	Ser	Trp	Ala	Arg	Met	Glu	Leu	Ser	Ile	Lys	Ser	Trp	Phe
385					390					395					400
Gly	Ala	Ala	Leu	Ala	Pro	Asp	Leu	Leu	Arg	Ser	Gly	Thr	Pro	Ser	Pro
				405					410					415	
His	Tyr	Glu	Ser	Ile	Leu	Arg	Leu	Ala	Ala	Ser	Gly	Pro	Pro	Gly	Gly
			420					425					430		
Arg	Gly	Ala	Val	Gly	Gly	Ser	Cys	Arg	Asp	Lys	Ile	Gln	Arg	Thr	Arg
		435					440					445			
Arg	Asp	Asn	Ala	Pro	Pro	Pro	Leu	Pro	Arg	Ala	Arg	Pro	His	Ser	Thr
	450						455				460				
Pro	Ala	Ala	Pro	Arg	Arg	Cys	Arg	Arg	His	Arg	Glu	Asp	Leu	Pro	Glu
465					470					475					480
Pro	Pro	His	Val	Asp	Ala	Ala	Asp	Arg	Gly	Pro	Glu	Pro	Cys	Ala	Gly
				485					490					495	
Arg	Pro	Ala	Thr	Tyr	Tyr	Thr	His	Met	Ala	Gly	Ala	Pro	Pro	Arg	Leu
			500					505						510	

Pro Pro Arg Asn Pro Ala Pro Pro Glu Gln Arg Pro Ala Ala Ala Ala
 515 520 525
 Arg Pro Leu Ala Ala Gln Arg Glu Ala Ala Gly Val Tyr Asp Ala Val
 530 535 540
 Arg Thr Trp Gly Pro Asp Ala Glu Ala Glu Pro Asp Gln Met Glu Asn
 545 550 555 560
 Thr Tyr Leu Leu Pro Asp Asp Asp Ala Ala Met Pro Ala Gly Val Gly
 565 570 575
 Leu Gly Ala Thr Pro Ala Ala Asp Thr Thr Ala Ala Ala Ala Trp Pro
 580 585 590
 Ala Glu Ser His Ala Pro Arg Ala Pro Ser Glu Asp Ala Asp Ser Ile
 595 600 605
 Tyr Glu Ser Val Gly Glu Asp Gly Gly Arg Val Tyr Glu Glu Ile Pro
 610 615 620
 Trp Val Arg Val Tyr Glu Asn Ile Cys Pro Arg Arg Arg Leu Ala Gly
 625 630 635 640
 Gly Ala Ala Leu Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala
 645 650 655
 Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro Arg
 660 665 670
 Arg Ala Thr Arg Ala Pro Asp Pro Gly Leu Ser Leu Ser Pro Met Pro
 675 680 685
 Ala Arg Pro Arg Thr Asn Ala Leu Ala Asn Asp Gly Pro Thr Asn Val
 690 695 700
 Ala Ala Leu Ser Ala Leu Leu Thr Lys Leu Lys Arg Gly Arg His Gln
 705 710 715 720
 Ser His

<210> 160
 <211> 318
 <212> PRT
 <213> HSV2

<400> 160
 Met Ile Thr Asp Cys Phe Glu Ala Asp Ile Ala Ile Pro Ser Gly Ile
 5 10 15
 Ser Arg Pro Asp Ala Ala Ala Leu Gln Arg Cys Glu Gly Arg Val Val
 20 25 30
 Phe Leu Pro Thr Ile Arg Arg Gln Leu Ala Leu Ala Asp Val Ala His
 35 40 45

Glu Ser Phe Val Ser Gly Gly Val Ser Pro Asp Thr Leu Gly Leu Leu
 50 55 60
 Leu Ala Tyr Arg Arg Arg Phe Pro Ala Val Ile Thr Arg Val Leu Pro
 65 70 75 80
 Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly Leu Thr His Ala Gly
 85 90 95
 Thr Val Asn Leu Arg Asn Thr Ser Pro Val Asp Leu Cys Asn Gly Asp
 100 105 110
 Pro Val Ser Leu Val Pro Pro Val Phe Glu Gly Gln Ala Thr Asp Val
 115 120 125
 Arg Leu Glu Ser Leu Asp Leu Thr Leu Arg Phe Pro Val Pro Leu Pro
 130 135 140
 Thr Pro Leu Ala Arg Glu Ile Val Ala Arg Leu Val Ala Arg Gly Ile
 145 150 155 160
 Arg Asp Leu Asn Pro Asp Pro Arg Thr Pro Gly Glu Leu Pro Asp Leu
 165 170 175
 Asn Val Leu Tyr Tyr Asn Gly Ala Arg Leu Ser Leu Val Ala Asp Val
 180 185 190
 Gln Gln Leu Ala Ser Val Asn Thr Glu Leu Arg Ser Leu Val Leu Asn
 195 200 205
 Met Val Tyr Ser Ile Thr Glu Gly Thr Thr Leu Ile Leu Thr Leu Ile
 210 215 220
 Pro Arg Leu Leu Ala Leu Ser Ala Gln Asp Gly Tyr Val Asn Ala Leu
 225 230 235 240
 Leu Gln Met Gln Ser Val Thr Arg Glu Ala Ala Gln Leu Ile His Pro
 245 250 255
 Glu Ala Pro Met Leu Met Gln Asp Gly Glu Arg Arg Leu Pro Leu Tyr
 260 265 270
 Glu Ala Leu Val Ala Trp Leu Ala His Ala Gly Gln Leu Gly Asp Ile
 275 280 285
 Leu Ala Leu Ala Pro Ala Val Arg Val Cys Thr Phe Asp Gly Ala Ala
 290 295 300
 Val Val Gln Ser Gly Asp Met Ala Pro Val Ile Arg Tyr Pro
 305 310 315

<210> 161
 <211> 825
 <212> PRT
 <213> HSV2

<400> 161

Met	Glu	Pro	Arg	Pro	Gly	Thr	Ser	Ser	Arg	Ala	Asp	Pro	Gly	Pro	Glu
				5					10					15	
Arg	Pro	Pro	Arg	Gln	Thr	Pro	Gly	Thr	Gln	Pro	Ala	Ala	Pro	His	Ala
			20					25					30		
Trp	Gly	Met	Leu	Asn	Asp	Met	Gln	Trp	Leu	Ala	Ser	Ser	Asp	Ser	Glu
		35					40					45			
Glu	Glu	Thr	Glu	Val	Gly	Ile	Ser	Asp	Asp	Asp	Leu	His	Arg	Asp	Ser
	50					55					60				
Thr	Ser	Glu	Ala	Gly	Ser	Thr	Asp	Thr	Glu	Met	Phe	Glu	Ala	Gly	Leu
65					70					75					80
Met	Asp	Ala	Ala	Thr	Pro	Pro	Ala	Arg	Pro	Pro	Ala	Glu	Arg	Gln	Gly
				85					90					95	
Ser	Pro	Thr	Pro	Ala	Asp	Ala	Gln	Gly	Ser	Cys	Gly	Gly	Gly	Pro	Val
			100					105					110		
Gly	Glu	Glu	Glu	Ala	Glu	Ala	Gly	Gly	Gly	Gly	Asp	Val	Cys	Ala	Val
		115					120					125			
Cys	Thr	Asp	Glu	Ile	Ala	Pro	Pro	Leu	Arg	Cys	Gln	Ser	Phe	Pro	Cys
	130					135					140				
Leu	His	Pro	Phe	Cys	Ile	Pro	Cys	Met	Lys	Thr	Trp	Ile	Pro	Leu	Arg
145					150					155					160
Asn	Thr	Cys	Pro	Leu	Cys	Asn	Thr	Pro	Val	Ala	Tyr	Leu	Ile	Val	Gly
				165					170					175	
Val	Thr	Ala	Ser	Gly	Ser	Phe	Ser	Thr	Ile	Pro	Ile	Val	Asn	Asp	Pro
			180					185					190		
Arg	Thr	Arg	Val	Glu	Ala	Glu	Ala	Ala	Val	Arg	Ala	Gly	Thr	Ala	Val
		195					200					205			
Asp	Phe	Ile	Trp	Thr	Gly	Asn	Pro	Arg	Thr	Ala	Pro	Arg	Ser	Leu	Ser
	210					215					220				
Leu	Gly	Gly	His	Thr	Val	Arg	Ala	Leu	Ser	Pro	Thr	Pro	Pro	Trp	Pro
225					230					235					240
Gly	Thr	Asp	Asp	Glu	Asp	Asp	Asp	Leu	Ala	Asp	Val	Asp	Tyr	Val	Pro
				245					250					255	
Pro	Ala	Pro	Arg	Arg	Ala	Pro	Arg	Arg	Gly	Gly	Gly	Gly	Ala	Gly	Ala
			260					265					270		
Thr	Arg	Gly	Thr	Ser	Gln	Pro	Ala	Ala	Thr	Arg	Pro	Ala	Pro	Pro	Gly
		275					280					285			
Ala	Pro	Arg	Ser	Ser	Ser	Ser	Gly	Gly	Ala	Pro	Leu	Arg	Ala	Gly	Val
	290					295					300				
Gly	Ser	Gly	Ser	Gly	Gly	Gly	Pro	Ala	Val	Ala	Ala	Val	Val	Pro	Arg

305		310		315		320									
Val	Ala	Ser	Leu	Pro	Pro	Ala	Ala	Gly	Gly	Gly	Arg	Ala	Gln	Ala	Arg
				325					330					335	
Arg	Val	Gly	Glu	Asp	Ala	Ala	Ala	Ala	Glu	Gly	Arg	Thr	Pro	Pro	Ala
			340					345					350		
Arg	Gln	Pro	Arg	Ala	Ala	Gln	Glu	Pro	Pro	Ile	Val	Ile	Ser	Asp	Ser
		355					360					365			
Pro	Pro	Pro	Ser	Pro	Arg	Arg	Pro	Ala	Gly	Pro	Gly	Pro	Leu	Ser	Phe
		370				375					380				
Val	Ser	Ser	Ser	Ser	Ala	Gln	Val	Ser	Ser	Gly	Pro	Gly	Gly	Gly	Gly
385					390					395					400
Leu	Pro	Gln	Ser	Ser	Gly	Arg	Ala	Ala	Arg	Pro	Arg	Ala	Ala	Val	Ala
				405					410					415	
Pro	Arg	Val	Arg	Ser	Pro	Pro	Arg	Ala	Ala	Ala	Ala	Pro	Val	Val	Ser
			420					425					430		
Ala	Ser	Ala	Asp	Ala	Ala	Gly	Pro	Ala	Pro	Pro	Ala	Val	Pro	Val	Asp
		435					440					445			
Ala	His	Arg	Ala	Pro	Arg	Ser	Arg	Met	Thr	Gln	Ala	Gln	Thr	Asp	Thr
	450					455					460				
Gln	Ala	Gln	Ser	Leu	Gly	Arg	Ala	Gly	Ala	Thr	Asp	Ala	Arg	Gly	Ser
465					470					475					480
Gly	Gly	Pro	Gly	Ala	Glu	Gly	Gly	Pro	Gly	Val	Pro	Arg	Gly	Thr	Asn
				485					490					495	
Thr	Pro	Gly	Ala	Ala	Pro	His	Ala	Ala	Glu	Gly	Ala	Ala	Ala	Arg	Pro
			500					505					510		
Arg	Lys	Arg	Arg	Gly	Ser	Asp	Ser	Gly	Pro	Ala	Ala	Ser	Ser	Ser	Ala
		515					520					525			
Ser	Ser	Ser	Ala	Ala	Pro	Arg	Ser	Pro	Leu	Ala	Pro	Gln	Gly	Val	Gly
		530				535					540				
Ala	Lys	Arg	Ala	Ala	Pro	Arg	Arg	Ala	Pro	Asp	Ser	Asp	Ser	Gly	Asp
545					550					555					560
Arg	Gly	His	Gly	Pro	Leu	Ala	Pro	Ala	Ser	Ala	Gly	Ala	Ala	Pro	Pro
				565				570						575	
Ser	Ala	Ser	Pro	Ser	Ser	Gln	Ala	Ala	Val	Ala	Ala	Ala	Ser	Ser	Ser
			580				585						590		
Ser	Ala	Ser	Ser	Ser	Ser	Ala	Ser	Ser	Ser	Ser	Ala	Ser	Ser	Ser	Ser
		595					600					605			
Ala	Ser	Ser	Ser	Ser	Ala	Ser	Ser	Ser	Ser	Ala	Ser	Ser	Ser	Ser	Ala
		610				615					620				

Ser Ser Ser Ala Gly Gly Ala Gly Gly Ser Val Ala Ser Ala Ser Gly
 625 630 635 640
 Ala Gly Glu Arg Arg Glu Thr Ser Leu Gly Pro Arg Ala Ala Ala Pro
 645 650 655
 Arg Gly Pro Arg Lys Cys Ala Arg Lys Thr Arg His Ala Glu Gly Gly
 660 665 670
 Pro Glu Pro Gly Ala Arg Asp Pro Ala Pro Gly Leu Thr Arg Tyr Leu
 675 680 685
 Pro Ile Ala Gly Val Ser Ser Val Val Ala Leu Ala Pro Tyr Val Asn
 690 695 700
 Lys Thr Val Thr Gly Asp Cys Leu Pro Val Leu Asp Met Glu Thr Gly
 705 710 715 720
 His Ile Gly Ala Tyr Val Val Leu Val Asp Gln Thr Gly Asn Val Ala
 725 730 735
 Asp Leu Leu Arg Ala Ala Ala Pro Ala Trp Ser Arg Arg Thr Leu Leu
 740 745 750
 Pro Glu His Ala Arg Asn Cys Val Arg Pro Pro Asp Tyr Pro Thr Pro
 755 760 765
 Pro Ala Ser Glu Trp Asn Ser Leu Trp Met Thr Pro Val Gly Asn Met
 770 775 780
 Leu Phe Asp Gln Gly Thr Leu Val Gly Ala Leu Asp Phe His Gly Leu
 785 790 795 800
 Arg Ser Arg His Pro Trp Ser Arg Glu Gln Gly Ala Pro Ala Pro Ala
 805 810 815
 Gly Asp Ala Pro Ala Gly His Gly Glu
 820 825

<210> 162

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 162

Ile Trp Thr Gly Asn Pro Arg Thr Ala Pro Arg Ser Leu Ser Leu
 1 5 10 15

<210> 163

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 163

Tyr	Met	Phe	Phe	Met	Arg	Pro	Ala	Asp	Pro	Ser	Arg	Pro	Ser	Thr
1				5					10					15

<210> 164

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 164

Val	Cys	Arg	Arg	Leu	Gly	Pro	Ala	Asp	Arg	Arg	Phe	Val	Ala	Leu
1				5					10					15

<210> 165

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 165

Gly	Pro	Ala	Asp	Arg	Arg	Phe	Val	Ala	Leu	Ser	Gly	Ser	Leu	Glu
1				5					10					15

<210> 166

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 166

Ser	Asp	Val	Leu	Gly	His	Leu	Thr	Arg	Leu	Ala	His	Leu	Trp	Glu
1				5					10					15

<210> 167

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 167

Gly His Met Thr Ile Ser Thr Ala Ala Gln Tyr Arg Asn Ala Val
1 5 10 15

<210> 168
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 168
Leu Asn Ala Trp Arg Gln Arg Leu Ala His Gly Arg Val Arg Trp
1 5 10 15

<210> 169
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 169
Gln Arg Leu Ala His Gly Arg Val Arg Trp Val Ala Glu Cys Gln
1 5 10 15

<210> 170
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 170
Asp Leu Val Ala Ile Gly Asp Arg Leu Val Phe Leu Glu Ala Leu
1 5 10 15

<210> 171
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 171
Gly Asp Arg Leu Val Phe Leu Glu Ala Leu Glu Arg Arg Ile Tyr
1 5 10 15

<210> 172
<211> 15

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 172
Leu Ala Tyr Arg Arg Arg Phe Pro Ala Val Ile Thr Arg Val Leu
1 5 10 15

<210> 173
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 173
Arg Arg Phe Pro Ala Val Ile Thr Arg Val Leu Pro Thr Arg Ile
1 5 10 15

<210> 174
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 174
Cys Ala Ile Ile His Ala Pro Ala Val Ser Gly Pro Gly Pro His
1 5 10 15

<210> 175
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 175
Pro Asn Gly Thr Arg Gly Phe Ala Pro Gly Ala Leu Arg Val Asp
1 5 10 15

<210> 176
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 176

Leu Arg Val Leu Arg Ala Ala Asp Gly Pro Glu Ala Cys Tyr Val
1 5 10 15

<210> 177

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 177

Asn Pro Arg Thr Ala Pro Arg Ser Leu
1 5

<210> 178

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic peptide

<400> 178

Gly Pro Ala Asp Arg Arg Phe Val Ala Leu
1 5 10

<210> 179

<211> 2100

<212> DNA

<213> HSV-2

<400> 179

atgcacgccca tgcgtcccag gttgcttctt ctttttggtc tttctggtct tccgggggaca 60
cgcgggcggt cggtgtccc cggaccaatt aatcccccca acagcgatgt tgttttcccg 120
ggagggtccc cgtgggtca atattgttat gcctatcccc ggttgacga tcccgggccc 180
ttgggttccg cggacgccg gcggcaagac ctgccccggc gcgtcgtccg tcacgagccc 240
ctgggcccgt cgttcctcac gggggggctg gttttgctgg cgccgccggt acgaggattt 300
ggcgaccca acgcaacgta tgcggcccgt gtgacgtact accggctcac ccgcgcctgc 360
cgtcagccca tcctccttcg gcagtatgga ggggtgtcgc gcggcgagcc gccgtcccca 420
aagacgtgcg ggtcgtacac gtacacgtac cagggcggcg ggccctccgac ccggtacgct 480
ctcgtaaatg cttccctgct ggtgccgatc tgggaccgcg ccgcggagac attcgagtac 540
cagatcgaac tcggcggcga gctgcacgtg ggtctgttgt gggtagaggt gggcggggag 600
ggccccggcc ccaccgcccc cccacaggcg gcgcgtgcg agggcgcccc gtgcgtcccc 660
ccggtccccc cgggccgccc gtggcgctcg gtgcccccg tatggtattc cgtccccaac 720
ccggggtttc gtggcctgcg tttccgggag cgctgtctgc cccacagac gcccgccgcc 780
cccagcgacc taccacgct cgcttttgct cccagagcc tgctggtggg gattacgggc 840
cgacggttta ttcggatggc acgacccacg gaagacgtcg ggtccctgcc gccccattgg 900
gccccgggg ccctagatga cgggtccgtac gcccccttc caccgcgcc gcggtttcga 960
cgcgccctgc ggacagaccc cgagggggtc gaccccgacg ttcgggcccc ccgaaccggg 1020
cggcgcctca tggccttgac cgaggacacg tcctccgatt cgctacgct cgctccggag 1080
aagacgcccc tcctgtgtc ggccaccgcc atggcaccct cagtcgaccc aagcgcgga 1140
ccgaccgcc ccgcaaccac tactcccccc gacgagatgg ccacacaagc cgcaacggtc 1200
gccgttacgc cggaggaaac ggcagtcgcc tccccgccg cgactgcac cgtggagtcg 1260

```

tcgccactcc ccgccgcggc ggcgggcaacg cccggggccg ggcacacgaa caccagcagc 1320
gcctccgcag cgaaaacgcc ccccaccaca ccagccccc cgaaccccc gccacgtct 1380
accacgcga cccccgcc cagactccg gggcccaaa caaccctcc cggaccgca 1440
ccccgggtc cgggtgggcg ctcgcgcgcg ccacggcgc attccccct caccgcctcg 1500
cccccgcta ccgcgcggg gccctcggc gccaacgttt cggtcgcgc gaccaccgc 1560
acgcccggaa cccggggcac cgccgtacc ccccaacgg acccaaagac gcaccacac 1620
ggaccgcgg acgtccccc cggctcgcca gccccccac ccccggaaca tcgcggcgga 1680
cccgaggagt ttgagggcgc cggggacggc gaacccccg aggacgacga cagcgccacc 1740
ggcctcgct tccgaactcc gaacccaac aaaccacccc ccgcgcgcc cgggccatc 1800
cgccccacgc tcccgcagg aattcttgg ccgctcgccc ccaacacgcc tcgcccccc 1860
gccaaagctc ccgctaagga catgccctcg gggccacac cccaacacat ccccctgtt 1920
tggttcctaa cggcctcccc tgctctagat atcctcttta tcatcagcac caccatccac 1980
acggcgcggt tcgtttgtct ggtcgccctg gcagcacaac tttggcgcg cggggcggg 2040
cgcaggcgat acgcgcacc gagcgtgcgt tacgtatgtc tgccaccga gcgggattag 2100

```

<210> 180
 <211> 471
 <212> DNA
 <213> Homo sapiens

```

<400> 180
atgcagcatc accaccatca ccacctgggt ctggtgaca cggtggtcgc gtgcgtggcc 60
ctggccgcgt ttgacggcgg gtcgacggcc cccgagggtg gcacgtacac ccccctgcgc 120
tacgcgtgcg tcttcgcgc gaccagccc ctgtacgcgc ggaccacccc cgccaaattt 180
tggcgggacg tgcgcgcgc cgcggaacac gtggaccttc gcccgcgctc ctgcgcgccc 240
cgggcgcccc tgagcgggac ggcagacccc gccttcctgc tcgaagacct ggcgcccttc 300
cccccgccc cctgaatag cgagtcgctg ctggggccgc gggtcgcgt cgtggacatc 360
atggcgcagt ttcggaact gtcctgggc gacgaggaga ccgcgcct cggggcgcac 420
gtgtccggga ggcgcgcgac cgggctgggc ggcccgcac gcccatagt a 471

```

<210> 181
 <211> 155
 <212> PRT
 <213> Homo sapiens

```

<400> 181
Met Gln His His His His His His Leu Gly Leu Ala Asp Thr Val Val
          5              10              15

Ala Cys Val Ala Leu Ala Ala Phe Asp Gly Gly Ser Thr Ala Pro Glu
          20              25              30

Val Gly Thr Tyr Thr Pro Leu Arg Tyr Ala Cys Val Leu Arg Ala Thr
          35              40              45

Gln Pro Leu Tyr Ala Arg Thr Thr Pro Ala Lys Phe Trp Ala Asp Val
          50              55              60

Arg Ala Ala Ala Glu His Val Asp Leu Arg Pro Ala Ser Ser Ala Pro
          65              70              75              80

Arg Ala Pro Val Ser Gly Thr Ala Asp Pro Ala Phe Leu Leu Glu Asp
          85              90              95

Leu Ala Ala Phe Pro Pro Ala Pro Leu Asn Ser Glu Ser Val Leu Gly
          100             105             110

```

Pro Arg Val Arg Val Val Asp Ile Met Ala Gln Phe Arg Lys Leu Leu
 115 120 125

Met Gly Asp Glu Glu Thr Ala Ala Leu Arg Ala His Val Ser Gly Arg
 130 135 140

Arg Ala Thr Gly Leu Gly Gly Pro Pro Arg Pro
 145 150 155

<210> 182
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 182
 ctgggtctgg ctgacacggt ggctgcgtgc gtg 33

<210> 183
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 183
 ccgttagaat tcactatggg cgtggcgggc c 31

<210> 184
 <211> 957
 <212> DNA
 <213> HSV2

<400> 184
 atgatcacgg attgtttcga agcagacatc gcgatcccct cgggtatctc gcgccccgat 60
 gccgcggcgc tgcagcgggtg cgaggggtcga gtggtctttc tgccgaccat ccgccgccag 120
 ctggcgctcg cggacgtggc gcacgaatcg ttctgtctccg gaggagttag tcccgcacag 180
 ttggggttgt tgctggcgta ccgcaggcgc ttccccgcgg taatcacgcg ggtgctgccc 240
 acgcgaatcg tcgcctgccc cgtggacctg gggctcacgc acgccggcac cgtcaatctc 300
 cgcaacacct cccccgtoga cctctgcaac ggggatcccg tcagcctcgt cccgcccgtc 360
 ttcgagggcc aggcgacgga cgtgcgcctg gagtcgctgg acctcacgct gcggtttccg 420
 gtcccgtctc caacgcccct ggcccgcgag atagtcgcgc ggctggctgc ccggggcacc 480
 cgggacctca accccgaccc ccggacgccc ggggagctcc ccgacctcaa cgtgctgtat 540
 tacaacgggg cccgtctctc gctcgtggcc gacgtccagc aactcgctc cgtaaacacc 600
 gagctgcggt cgctcgctct caacatggtc tactccataa ccgaaggac caccctcatc 660
 ctcacgtcga tcccccggt gctcgcgctg agcgcgccag acggatacgt gaacgcgctc 720

ctgcagatgc	agagcgtcac	gcgagaagcc	gcccagctca	tccaccccca	agcccccatg	780
ctgatgcagg	acggcgaaacg	caggctgccc	ctttacgagg	cgctgggtcgc	ctggctggcg	840
cacgcggggc	aactcgggga	catcctggcc	ctggcccccg	cggttcgggt	gtgtacgttc	900
gacggcgccg	ccgttgtgca	atccggcgac	atggcccccg	ttatccgcta	cccctga	957

<210> 185
 <211> 1260
 <212> DNA
 <213> HSV2

<400> 185						
atgcctcggg	tctcctcttc	ctgtccctcc	tcctttctcc	cagccagggg	gaggaggggc	60
gggcgtggcg	ggcaggtgtg	cgggcggggg	gggcgcggcg	gcgggggtgg	gcacggggcg	120
aagtgcgggt	gcagcctcgc	ggtcttctct	tctcctctcc	ccttctctcc	acccgtcccc	180
gggggcagag	ggcgtgcatg	cgttgtgatt	caaccgcctc	cgcccccgcc	ccactttccc	240
ccctctctat	caaagtcccc	tggccctcgg	cttcgcgcgc	gtgggtgcggc	tgaccccccc	300
cctcctccct	ccccgagcca	ggcgccctcc	cactcctgcc	caccaccccc	aggggtctggc	360
cggccagacg	tgcgtgctct	gcacgatcgg	gccccctccc	ctgtcaacac	ggacacactc	420
tttttttacc	cgccagccag	cccgccccac	caccaagaca	gggagccaga	acgagggcgg	480
gccccggctc	tggtctatga	ttaaagacca	caggcctcgg	gggtgggggc	ggcttctcgt	540
gcccccccc	cctcctcctc	ctcccttccc	ccccatcccc	ggccccctcg	cgcgggggag	600
ctgcatcaaa	ggccaacaac	aaagtgtgtc	aaaagcatca	caaaaacttta	ttgtaaaatt	660
tttataaata	taaagtthtt	ttttctctca	agttttcaac	aaggccagaa	agtccataac	720
aaaatgctgg	tgtgtgttgc	tggtcggggc	cgtgtccgtc	cccccccccc	actcccaccc	780
ccacttcctg	tctcctcccc	gtctttcccc	ccccccacct	ccccctgccc	ccgagggcgc	840
tcggccgggt	gtccgggtgg	gggcggcttc	cttcggggcag	caagccgagt	gttagctccc	900
cctactcccc	gtggcccgcg	ggggcgctgc	cggccggcgc	gggcgcgcgc	tgctcccag	960
accacgggtg	gcgcgaccgg	aggccgtgga	agtcacagcg	gcccaccagg	gtgccctgg	1020
caaagagcat	ggtgcccacc	gggtcatccc	agaggtgtgt	ccactccgac	gcggggggcg	1080
tcgggtagtc	ggggggcctc	acgcagttgc	gcgcgtgctc	ggggagcagg	gtgcggcggc	1140
tccacgcggg	ggccgcggcc	cgacgcaggt	ccgccacgtt	ccccgtctgg	tccacgagga	1200
ccacgtaggc	ccctatgtgg	cccgtctcca	tgtccaggac	gggcaggcag	tcccccgta	1260

<210> 186
 <211> 1473
 <212> DNA
 <213> HSV2

<400> 186						
atggacacct	tggtcgacga	tctgtttgcg	gacgcggacg	gggtttcgcc	accgcccccc	60
aggccagccg	gggggtcccaa	gaacacccca	gcgcgccctc	cgctgtacgc	caccgggtcg	120
ctgagtcagg	cccagctgat	gccctcgccg	cccatgccc	tccccccgc	ggccctgttt	180
aaccgtctcc	tcgacgatct	gggttccagc	gcgggtcccc	cgctgtgtac	catgctagat	240
acctggaacg	aggacctgtt	ctctgggttc	ccgaccaacg	ccgacatgta	ccgggagtg	300
aagttttctg	cgacgctgcc	cagcgacgtg	atcgactggg	gggatgcgca	cgtccccgag	360
cgctccccga	tcgacattcg	cgcccacggc	gacgtggcgt	tccccaccct	gcccgcaccc	420
cgcgacgagc	tgcccttcgta	ctacgaggcc	atggcgcagt	ttttccgcgg	tgagctgcgg	480
gcgcgggagg	agagctaccg	gaccgtgttg	gcaaattttt	gctcggccct	gtaccgggtac	540
ctgcgcggca	gcgttcggca	gctacaccgc	caggcacaca	tgccggggcg	caaccgcgac	600
ctgcgcggaga	tgctgcgcac	cacgatcgcg	gacaggtact	accgcgagac	cgcgcgccct	660
gcgcgcgtcc	tggtttctgca	tctatacctc	tttctgagcc	gcgagatcct	atgggcccgc	720
tacgccgagc	agatgatgcg	gcccgatctg	ttcgacggcc	tctgtctcga	cctggagagc	780
tggcgccagt	tggtgtgtct	gtttcagccc	ctgatgttta	tcaacggatc	gctcaccgtg	840
cggggagttc	ccgtggaggc	ccggcgactg	cgggagctaa	accacattcg	cgagcacctg	900
aacctcccgc	tggtgcgaag	tgcggcggcg	gaggaaaccc	gggcgcccct	cacgacccc	960
cccgtcctgc	agggcaacca	ggcccgtccc	tctgggtact	ttatgctgct	gatccggggc	1020
aagttggact	cgtactccag	cgtcgcgacc	tcggagggcg	agtccgtcat	gcgggagcac	1080
gcgtatagcc	gcggggcgac	cagaaacaat	tacggatcga	caatcgaggg	cctgctcgac	1140

ctccccgacg	acgatgacgc	tccgtcggag	gcggggctgg	tggcgccgcg	catgtcgttt	1200
ctctccgcgg	gacaacgccc	ccgcagactg	tccaccaccg	ccccattac	cgacgtcagc	1260
ctgggagacg	aactccgcct	ggacggcgag	gaggtggata	tgacgcccgc	cgacgccctg	1320
gacgacttcg	acttgagat	gctgggggac	gtggagtccc	cctccccggg	aatgaccac	1380
gaccccgctc	cgtatggggc	tttgacgtg	gacgattttg	agtttgaaca	gatgtttacc	1440
gatgccatgg	gcattgacga	ctttgggggg	tag			1473

<210> 187
 <211> 1479
 <212> DNA
 <213> HSV2

<400> 187						
atgggtctgt	ttggcatgat	gaagtttgcc	cagactcacc	atctggtgaa	gcgcgggggc	60
ctccggggccc	cggagggcta	ctttaccccc	atcgccgtgg	acctgtggaa	tgtcatgtat	120
accctggtgg	ttaaatatca	gcgcgctac	ccaagttacg	accgcgaggc	aatcacgcta	180
cactgtctct	gtagtatggt	acgggtgttt	acccaaaagt	ccctgttccc	catcttcgtg	240
accgatcgcg	gggtcgagtg	taccgagccg	gttgtgttcg	gggccaaggc	gatcctggcc	300
cgcacgacgg	cccagtgcgg	cacggacgag	gagggcagtg	acgtagacgc	ctcgccgcgg	360
ccttccccca	tcaccgactc	caggcccagt	ttcgcccttt	ccaacatgcg	ccgcgcgggg	420
cacgccttcg	ccccggggga	ccggggaacg	cgggccgcgg	gcccaggccc	ggcgggcccc	480
tggggcgcg	cctcgaagcc	ggccctgcgc	ctggctcacc	tgttctgtat	ccgcgttctg	540
cgggcgctgg	ggtacgccta	catcaactcg	ggtcagctgg	aggccgacga	cgcctgcgcg	600
aacctctatc	ataccaacac	ggtcgcgtac	gtgcatacca	cggataaccga	tctcctgctg	660
atgggctgcg	atatcgtggt	ggacatcagc	accggctaca	ttccgacgat	tactgcccgc	720
gacctgctgc	agtacttcaa	gatgagttac	ccgcagttcc	tggcgctggt	cgtccgctgc	780
cacacagacc	tgcaccccaa	taacacctac	gcgtccgtcg	aggacgtgct	gcgcgagtgt	840
cactggaccc	ccccgagccg	atcccagccc	cgcggggcgg	cccgcgggga	gcgcgccaac	900
tgcgctcccc	tggagagcat	gcctacgtcg	accgcggccc	cggtcggcct	cgagacgcgc	960
atctcgtgga	ccgaaattct	ggcccaacag	atcgcgggcg	aggacgacta	cgaagaagac	1020
ccccccctcc	agcccccgga	cgtcgcgggt	gggcccgcgc	acggcgcccc	gtcgtcctcc	1080
tggagatac	tcaccccgcc	cgagctcgtg	caggtcccca	acgcgcagcg	ggtcgcggaa	1140
caccgcggct	atgtcgcggg	acgtcgcgcg	cacgtcatcc	acgacgcccc	ggaggccctg	1200
gactggctgc	ccgatccgat	gaccatcgcc	gagctggtgg	agcacagata	cgtcaagtac	1260
gtcatatcgc	ttatcagccc	caaggagcgg	ggaccctgga	ctcttctaaa	aagactgccc	1320
atctatcagg	acctccgcga	cgaagattta	gcgcgctcca	tctgtactcg	gcataatcacc	1380
gccccggaca	tcgcccagccg	gtttctggcg	cagctgtggg	cccacgcgcc	cccgcgcgcg	1440
ttttacaagg	acgtcctggc	taaattctgg	gacgagtag			1479

<210> 188
 <211> 3429
 <212> DNA
 <213> HSV2

<400> 188						
atggccaacc	gccctgccgc	atccgcctcc	gccggagcgc	ggtctccgtc	cgaacgacag	60
gaaccccggg	agcccgaggt	cgccccccct	ggcgccgacc	acgtgttttg	caggaaagtc	120
agcggcgtga	tggtgctttc	cagcgatccc	ccgggccccg	cggcctaccg	cattagcgac	180
agcagctttg	ttcaatgcgg	ctccaactgc	agtatgataa	tcgacggaga	cgtggcgcg	240
ggtcatttgc	gtgacctcga	gggcgctacg	tccaccggcg	ccttcgtcgc	gatctcaaac	300
tgcgcagccg	gcggggatgg	ccgaacgcgc	gtcgtggcgc	tcggcggaac	ctcgggcccc	360
tccgcgacta	catccgtggg	gacccagacg	tccggggagt	tcctccacgg	gaacccaagg	420
acccccgaac	cccaaggacc	ccaggtgtgc	ccccgccccc	ctcctccccc	ctttccatgg	480
ggccacgagt	gctgcgcccc	tcgcgatgcc	agggcgggcg	ccgagaagga	cgtcggggcc	540
gcggagtcac	ggtcagacgg	cccgtcgtcc	gactccgaaa	cggaggactc	ggactcctcg	600
gacgaggata	cgggttcgga	gacgctgtct	cgatcctctt	cgatctgggc	cgcagggggc	660
actgacgacg	atgacagcga	ctccgactcg	cggtcggacg	actccgtgca	gccccacggt	720
gtcgttcgtc	gcagatggag	cgacggcccc	gcccccggtg	cctttcccaa	gccccggcgc	780

```

ccccgcgact cccccggaaa ccccgccctg ggccgcccga cccggccggg ctccgcgacg 840
gacccgcgcg cgtcggccga ctccgattcc gcggcccaag ccggccgacc ccaggcggac 900
gtggcgccgg ttctggacag ccagcccact gtgggaacgg accccggcta ccagtcctcc 960
ctagaactca cccccagaaa cgcggaggcg gtggcgccgt ttctggggga cgcgcgcgac 1020
cgcgagcccg cgtcatgct ggagtacttc cgcctcacgg aggacgactt tgggctcctg 1080
gtgccccac gaaccttcgg cagcgcccc cgcctcacgg aggacgactt tgggctcctg 1140
aactacgcgc tcgtgagat gcgacgcctg tgcttgacc ttccccgggt ccccccaac 1200
gcatacacgc cctatcatct gagggagtat gcgacgcggc tgggtaacgg gttcaaacc 1260
ctgggtgcggc ggtccgccc cctgtatcgc atcctggggg ttctggtcca cctgcgac 1320
cgtacccggg aggcctcctt tgaggaatgg atgcgctcca aggaggtgga cctggacttc 1380
gggctgacgg aaaggtctcg cgaacacgag gccagctaa tgatcctggc ccaggccctg 1440
aaccctacg actgtctgat ccacagcacc ccgaacacgc tcgtcgagcg ggggctgcag 1500
tcggcgctga agtacgaaga gttttacctc aagcgtctcg gcgggacta catggagtcc 1560
gtcttccaga tgtacacccg catcgccggg tttctggcgt gccgggcgac ccgcgcatg 1620
cgccacatcg ccttggggcg acaggggtcg tgggggaaa tgttcaagtt cttttccac 1680
cgctctacg accaccagat cgtgcccgtc acccccgcca tgctgaacct cggaaccgc 1740
aactactaca cgtccagctg ctacctgga aaccccaggg ccaccactaa ccaggccacc 1800
actccggcca tcacggcaa cgtgagcgcc atcctcgccc gcaacggggg catcgggctg 1860
tgcatgcagg cgttcaacga cgcagcccc ggcacgcca gcatcatgcc ggccctgaag 1920
gtcctcgact ccttggggc ggcgcacaac aaacagagca cgcgccccac cggggcgtgc 1980
gtgtacctgg aacctggca cagcgacgtt cgggcccgtc tcagaatgaa gggcgctcctc 2040
gccggcgagg aggcacagcg ctgcgacaac atcttcagcg cctctggat gccggacctg 2100
ttcttcaagc gcctgatccg ccacctcgac ggcgagaaaa acgtcacctg gtccctgttc 2160
gaccgggaca ccagcatgtc gtcgcgcgac tttcacggcg aggagttcga gaagctgtac 2220
gagcaccctg aggcctagg gttcggcgaa acgatcccca tccaggacct ggcgtacgcc 2280
atcgtgcgca gcgcggccac caccggaagc ccttcatca tgtttaagga cgcggtaaac 2340
cgccactaca tctacgacac gcaagggcg gccatcgccg gctccaacct ctgcaccgag 2400
atcgtccacc cggcctccaa gcgatccagt ggggtctgca acctgggaag cgtgaatctg 2460
gccgatgcg tctccaggca gacgtttgac tttggcgggc tccgcgacgc cgtgcaggcg 2520
tgctgtctga tgggtaacat catgatcgac agcacgctac aaccacgcc ccagtgcacc 2580
cgcggcaacg acaacctgcg gtccatgggc attggcatgc agggcctgca cacggcgtgc 2640
ctcaagatgg gcctggatct ggagtcggcc gagttccggg acctgaacac acacatcgcc 2700
gaggtgatgc tgctcgcgcc catgaagacc agtaacgcgc tgtgcgttcg cggggcgcg 2760
cccttcagcc actttaagcg cagcatgtac cgggcccggc gctttcactg ggagcgcttt 2820
tcgaacgcca gcccgcggtg cgagggcgag tgggagatgc tadgccagag catgatgaaa 2880
cacggcctgc gcaacagcca gttcatcgcg ctcatgcccc ccgcccctc ggccagatc 2940
tcggacgtca gcgagggctt tgccccctg ttcaccaacc tgttcagcaa ggtgaccagg 3000
gacggcgaga cgtgcgccc caacacgctc ttgctgaagg aactcgagcg cacgttcggc 3060
gggaagcgcg tcttgagcg gatggacgg ctcgagggcca agcagtggtc tgtggccag 3120
gccctgcctt gcctggaccc cgcaccccc ctccggcggt tcaagacggc cttcgactac 3180
gaccaggaac tgctgatoga cctgtgtgca gaccgcgccc cctatgttga tcacagccaa 3240
tccatgactc tgtatgtcac agagaaggcg gacgggacgc tccccgcctc caccctggtc 3300
cgcttctcg tccacgcata taagcgcggc ctgaagacgg ggatgtacta ctgcaaggtt 3360
cgcaaggcga ccaacagcgg ggtgttcgcc ggcgacgaca acatcgtctg cacaagctgc 3420
gcgtgttaa 3429

```

<210> 189

<211> 3345

<212> DNA

<213> HSV2

<400> 189

```

atgtcggaca ggcggtcca ggtcccggcg cccgcccga tgaccccccc gtccgctcca 60
ccccaaaacg gacggttgca ggtcctgttg gggagcctga ccaacctccg ccggccgccc 120
tccccctcga gcgagccggc gggcagcgcg gacgagccgg cgtttctgtc cgcggccaaa 180
ctgcacgccg caacggcggc gtttctgttg tcggcgcgcg cggtcggccc ggcgagggcg 240
cgggcgtgct ggcacccgct cctggagcaa ctgtgcgcct tgcacggggc ccacgggctt 300
ccggagacgg cgtctcttggc cgagaacctc cccgggctgc tgggtgcaccg catggcggtg 360

```

```

gcgctccccg agacccccga ggcggccttc cgggagatgg acgtcatcaa ggacacggtc 420
ctggcgatca ccggctccga caccgcccac gcgctggagg ccgccgggct gcgcaccacg 480
gcggccctgg ggccgggtgcg cgtccgccag tgtgccgtgg agtggatcga ccgctggcgg 540
accgtcaccg agagctgcct ggccatgaac ccgcggacct ccctggaggc cctcggggag 600
atgtcgttta agatgtcccc ggttccccctg gggcaaccgc gcgccaacct gaccacccccg 660
gcgtacagcc tgctcttccc gtccccgatc gtgcaagagg gtctccgggt cttggccctg 720
gtcagcaatt ggggtgacgt gttctcggcg cacctacagc gcatcgacga cgccgccctg 780
acccccctca cgccggccct gtttacgtg gcgctgggtg acgactacct gacgaccccc 840
gatcggggcg cgggtggtccc gccgccgctg ctggcccagt ttcagcacac ggtgcgggag 900
atcgaccggg ctatcatgat cccgccctc gagggcacca aaatggttcg cagtgcgcgag 960
gaggtgcgcg tgtcgaccgc cctaagccgc gtcagcccgc ggtcggcggt cgccgccccg 1020
gggacgctga tggcgcgcggt gcggacggac gcggccgtct tcgaccccca cgtgccgttt 1080
ctgagtgcgt cggcgctggc catcttccgt ccggccgtca ccggcctgct gcagctaggg 1140
gagccgccct ccgccggggc ccagcaacgg ctccctcgcg tgctgcagca gacgtgggtc 1200
ttggtccaga acagcaactc gccctccgtg gtgatcaaca cgcttaccca cgccggggtc 1260
acgccgcgcg actgcaccca gtacatatcg gccctggagg ggtttttggt ggccgggggtg 1320
cccgcgcgga cccccccgg ccacgggctc agcgagatcc agcagctctt tgggtgcatt 1380
gcccttgccg gggcgaaagt gttcgggctc gcgcgggagt acgggcacta cgccggctac 1440
gtgaaaacgt tccggcggtt ccaggggccc agcgagcaca cgcacgggag gctctgtgag 1500
gcggtcggcc tgtcggggcg cgtcctcagc cagacgctgg cgcgcacatc gggcccggcc 1560
gtgccgacgg aacacctggc gagcctgcgg cgcacgctcg tgggggagtt tgagacggcc 1620
gagcgccgct tcagcgccgg ccagcccagc ctccctgcgc agacggcgct catctggctc 1680
gacgtgtacg gccagaccca ctgggacctg acccccacca ccccggccac gccgctgtcc 1740
gcgctgtccc ccgtcgggac gcctagccac gccccctcgg ttacactggc cgccggccacc 1800
aagatccgct tcccgccctt ggagggcatc caccacaacg tccctcgcca cccgggggtc 1860
gtcccttacg tgcctggcct ggtggtcggg gacgcgtga gggccacgtg caacgctgcc 1920
tacctgcccc tccccatcga gtttgccctg cgggtgctgg cctgggcgcg cgacttcggc 1980
ctgggctatc tccccaccgt cgaggggcac cgcacaaaat tgggcgcgct gatcacctc 2040
ctcgaaccgg ccacccgggc cggcgtcggg cccacgatgc agatggcca caacatcgag 2100
cagctgtgcg gcgagctgta cgtgatcgcc aggggggccc tcgagcagct gcggcccgcg 2160
gtccagctgc cccccccca gcctcccgag gtgggcagca gcctgctgct gatcagcatg 2220
tacgccctgg cggcccgggg ggtgctgcag gagctcgccg agcgcgcgca cccctggtc 2280
cgccagctgg aggacgcgat cgtgctgctg cgggtgcaca tgcgcacgct cgccgccttc 2340
tttgagtgcc ggttcgagag cgacgggcac cgctgtacg ccgtggtcgc agacggccac 2400
gagcgctggg ggcctggcg ccccgaggcc atgggcgacg cggtagacca gtactgcggc 2460
atgtatcacg acgccaagcg cgcgtggtc gcgtccctcg cggggtgctg gtccgtcgtc 2520
accgacga cggcgcacct cggggtatgc gacgagctgg cggccaggt ctgcacgag 2580
ggtaacgtgc tggccgtcgt gcggcgggaa atccacgggt tccctggccat cgtgtcgggc 2640
atccacgcca gggcgctcaa gttgatgtcg ggggaccagg tccccgggtt ctgttacatg 2700
agtacgttcc tggcgcgctg gcggcgctc tcggccggtt atcaggccgc acgcgcggcc 2760
acgggccccg agcgggtggc cgagttcgtt caggaaactc acgacacgtg gaaggccctg 2820
cagacggagc gcgcgctggt cgtggcgcg ttcgccagct cggccgatca gcgcaccgca 2880
gccattcagg aggtgatggc ccacgcgacc gaggacgcgc ccccagccc ggccgcagac 2940
ctcgtcgtgc tcacgaaccg tcatgacctg ggggcctggg gggactacag cctggggccc 3000
ctgggtcagc cgacggtggt tccggactcc gtggacctgt ctcgccaggg gctggccgct 3060
acgctgagca tggactggct attaataaac gagctcctgc aggtcaccca cgcgctgttt 3120
cgcgcctcgg cgtttcggcc ttccgccggc ccgggggccc cgggggacct ggaggcccaa 3180
gatgccggcg gtagcaccac cgaaccacg acaccgggc cacaggacac gcaggccgcg 3240
gcgcgctcga cgcgcgggc gggccgcgag acggtccctt ggcccaacac ccccgtaggag 3300
gacgacgaga tgacgccgca ggagacacca ccggtacacc cgtag 3345

```

<210> 190

<211> 9369

<212> DNA

<213> HSV2

<400> 190

atgattccgg ccgcgctacc tcacccacg atgaagcgcc agggggatcg cgatatcgtg 60

gtcaaccggtg	ttcggaaacca	gttcgcgacc	gatctggagc	cggggggggtc	ggtatcgtgc	120
atgcgctcgt	ccctgtcgtt	tctcagtcct	cttttcgacg	tggggccccg	cgacgtgctg	180
tccgcggagg	ccatcgaggg	atgtctggtg	gaggggggcg	agtggacgcg	cgcgcccgcg	240
ggctccgggc	cgccgcgcat	gtgttccatt	atcgagctcc	ccaacttcct	ggagtaccog	300
gcggcgcgcg	gcggcctgcg	ttgcgtgttt	tcgcgggtgt	acggcgaggt	gggcttcttc	360
ggggagccca	ggcggggct	gctggagacg	caatgccccg	cgcacacgtt	cttcgcgggc	420
ccctgggcca	tgcggccgct	gtcctacacg	ctcctgacca	tcggccccct	gggcatgggg	480
ctctaccggg	acggcgacac	ggcctatctg	tttgaccgcg	acggcctgcc	cgcggaacg	540
cccgctttta	tcgccaaggt	gcgggcgggc	gacgtgtacc	cgtacctgac	gtattacgcc	600
cacgaccgcc	ccaaggtgcg	gtgggcgggc	gccatggtgt	tttttgtgcc	ctccggcccg	660
ggggccgtgg	cgccggcgga	cctgaccgcc	gcggccctgc	acctctacgg	ggccagcgag	720
acttacctgc	aggacgaacc	gttcgtcgag	cggcgcgtgg	ccatcaccca	ccccctgcgg	780
ggcgagatcg	gggggctcgg	ggcgctcttt	gtcgggtgtg	tccccagggg	ggacggagaa	840
ggctcggggc	cggtcgtccc	cgcgctcccc	gcgcccaccc	acgtccagac	ccctggggcc	900
gaccggcccc	ccgagggccc	ccgggggggc	tcggggcccac	ccgacacacc	gcaggcgggg	960
cacccaaacc	gccccccgga	tgatgtgtgg	gcagcagccc	tggaggggac	cccgcccgcg	1020
aagccctcgg	ccccggacgc	cgcgccctcg	gggtccaccgc	acgccgcccc	gcccccgag	1080
acaccgcgcg	gggacgcggc	cgaagaagcc	gaggacctgc	gggtcctgga	ggtgggggca	1140
gtgcgggtgg	gtcgccaccg	ggcgcgctat	tctaccggct	tgcccaagcg	ccgcggccc	1200
acctggaccc	caccgtccag	cgtggaggat	ctgacctcgg	gcgagaggcc	cgcccccaag	1260
gccccgcccc	ccaaggccaa	gaagaagtcc	gcgcccaga	agaaggcccc	cgtcgccgcc	1320
gaggtccccc	cctcctcccc	gacgcccata	gcgcgcacgg	tccctcccgc	accggatacc	1380
cccccgagct	ccgggcaggg	cggcggcgac	gacgggcccc	cgagcccgtc	gtcgccgtcc	1440
gtgctcgaga	cgctgggggc	ccgcgccccc	ccggagcccc	ccggcgccga	cctggcgag	1500
ctgttcgagg	tccaccccaa	tgtggccgcc	acggcggtcc	ggctcgcggc	gcgcgatgcc	1560
gccctggccc	gcgaggtggc	cgcgtgttcg	cagctcacca	tcaacgccct	gcggtcgccc	1620
taccgcgcgc	accgggggct	gttggagctc	tcgctcatct	tcttctttga	gcgcgtcctc	1680
gccttttttg	tcgagaacgg	ggcccgcacg	catacgaggg	ccgggggtggc	cgggcccgcc	1740
gcggccctgc	tggacttcac	cctgcgcgat	ctgcccgcga	aaacggccgt	cggggacttc	1800
ctggcgctcca	cgcgcatgag	cctggcgggc	gtggccggcc	accggcccct	catccagcac	1860
gtgctggacg	aaaattcgca	gatcgggcgc	ctggccctcg	ctaagctggt	cctggtcgcc	1920
agggacgtta	ttcgggagac	ggacgccttt	tacggggacc	tcgcggacct	ggatctgcag	1980
ctgcgcgcgc	ccccgcgggc	caatctgtac	gcacgcctcg	gagagtggct	cctggagcgc	2040
tcgcgcgcgc	accggaacac	cctgttcgcc	cccgccaccc	cgacgcaccc	ggagccgctt	2100
ctccatcgga	tccaggcgct	ggcccagttt	gcccggggcg	aagagatgcg	ggtcagggcg	2160
gagggccgcg	agatgcgcga	ggccttagac	gcccctcgcc	gcggggtcga	ctcggtctcg	2220
cagcgcgccg	ggcccctgac	cgtgatgccc	gtgcccgcgc	ccccgggcgc	cgcgggcagg	2280
gccccgtgcc	cgcccgccct	gggcccgcag	gccatccagg	cgcggtgga	ggacgtgcgg	2340
atccaggccc	gcccggcgat	cgagagcgcg	gtcaaggagt	acttccaccg	gggagccgta	2400
tacagcgcca	aggccctgca	ggccagcgac	agccacgact	gtcggtttca	cgtggcctcg	2460
gccgcgggtc	tgcccatggt	ccagttgtcg	gaatcgctac	cggcctttga	ccaacacacg	2520
cggaacgtcg	ctcagcgcg	cgcgctcccc	ccccgcgcgc	ccctcgccac	cagccccag	2580
gccatcctgc	tgcgggacct	gtccagcgcg	ggccagcccc	tagatgcccc	cgaggacctg	2640
gcggcctggc	tctcggtcct	gaccgacgcc	gcgaccaggg	ggctcatcga	acgcaagccc	2700
ctggaagagc	tggcgcgag	catccacggc	atcaacgacc	agcaggcgcg	gcgcagctcg	2760
gggctggccg	aactgcagcg	cttcgacgcc	ctcgacgcgg	cgctcgcgca	gcagctggac	2820
agcgacgcgg	ccttcgtccc	ggcgaccggc	cccgcgccgt	acgtcgacgg	ggcggggctg	2880
tcccccgagg	ccacgcgcat	ggccgaagac	gcgctgcggc	aggcgagggc	catggaggcg	2940
gccaaagtga	cggcgagctg	cgctccggag	gcgcgctccc	ggcttcggga	gcgcgcccac	3000
gcccgtggag	cgatgctcaa	cgacgcgcgg	gagcgcgcca	aggtggccca	cgacgcccgc	3060
gagaagtctc	tgacaaact	ccaaggggct	ctcgggcccc	tcccggactt	tgtggggcta	3120
aaggcgtgtc	cggccgtgct	ggcgaccctg	cgggcctccc	tgcccgcggg	gtggaccgac	3180
ctggccgatg	ccgtccgggg	cccccccccg	gaagtcaagg	cggcgctgcg	ggcggaacctg	3240
tgggggctgc	tggggcagta	ccgggagggc	ctggaacacc	cgaccccgga	taccgcgacg	3300
gctctggccg	gcctgcatcc	cgccttcgtg	gtggtgctga	agaccctgtt	cgccgacgcc	3360
cccgagacgc	cggctcctcgt	gcagttcttt	tcgatcacg	cccccaaat	cgccaaagcc	3420
gtgtcaaacc	ccatcaacgc	cgggagcgcc	gcgctcgcca	cggcaagtcc	cgcggcgacg	3480
gtggacgcgg	ccgtgcgggc	gcatggcgcc	ctggcgagac	cgggtgtcggc	cctggggcgcg	3540

gcccgcgcgc	acccggcctc	gcccctgtcc	ttcctcgcgc	tgctggccga	tagcgccgcg	3600
gggtatgtaa	aggcgactcg	gctggccctg	gaggcgccgg	gggccatcga	cgagctcacg	3660
accctggggg	cggcgccgc	cgacctggtg	gtccaggcgc	gccgggcctg	cgcccagccg	3720
gagggggacc	atgccgccct	catcgatgcc	gcggcccgcg	cgacgaccgc	tgctcgggaa	3780
agcctcgcgc	ggcacgagcg	tgggttcggg	ggcctgctgc	acgccgaagg	gacggccggg	3840
gaccattccc	ccagcgggcg	cgccctgcag	gagctgggca	aggtcatcgg	cgccacgcga	3900
cgccgcgcgc	acgaacttga	ggccgcgcgc	gcccagctga	cgccgaagat	ggcgcccag	3960
cgccgcgcgc	gcagcagcga	gcgctgggcc	gcccgcgtcg	aggcggcgct	ggaccgcgtc	4020
gagaaccgcg	ccgagttcga	cgtggctcag	ctgcgcgcgc	tgaggcgct	ggcgggcacg	4080
cacggctaca	acccccggga	cttcgcgaag	cgcgccgagc	aggcgctcgc	ggccaacgcc	4140
gagggcggtc	ccctcgcctc	ggatacggcc	ttcgcgttta	acccctacac	ccccgagaac	4200
cagcgccacc	ccatgctccc	cccgcctcgc	gccatccacc	gcctcggctg	gagcgccggc	4260
ttccaagccg	ccgccgaaac	ctacgcgcgc	atgtttcggg	tcgacgccga	gcccctggcg	4320
cggtgctgc	gcatcgccga	ggggtcctcg	gaaatggccc	aggcgggcga	cggttttatc	4380
gactaccacg	aggccgtggg	gcgcctggcg	gacgacatga	cgagcgtgcc	cggcctgcgc	4440
cggtacgtgc	cgttcttcca	acatggctac	gccgactacg	tggagctccg	cgatcgcccg	4500
gacgccatcc	gggcccagct	gcacccggcg	ctcggcgccg	tgccctgga	cctggcgccg	4560
gcccgcgagc	agatatccgc	ggcccgggaat	gacccggagg	ccaccgcgga	gctgggtccg	4620
accggggtca	ccctgccgtg	tcccagcgag	gacgcgctgg	tcgctgcgc	ggcggcgttg	4680
gagcgcggtg	accagagccc	cgtgaagaac	acggcgctacg	ccgagtatgt	cgcttcgtg	4740
acccggcagg	acacggccga	gaccaaggac	gcagtgggtg	gcgcaaaca	gcagcgccgc	4800
gagggccacc	agcgggtcat	ggcgggcctg	cgggaggcgc	tggccgcgcg	cgagcgccgc	4860
gcccagatcg	aggccgaggg	cctggccaac	ctcaagacca	tgctgaagg	ggtcgccgtc	4920
ccggcgaccg	tgjccaagac	gctcgaccag	gcgcgctcgc	tggcgagat	cgccgaccag	4980
gtcgaggttc	tggttgatca	gaccgagaaa	acgcgcgagc	tcgacgtgcc	cgcggttatc	5040
tggttagagc	acgccagcg	caccttcgag	acgcacccgc	taagcgcgc	ccgcggcgac	5100
ggccccggcc	ccctggcgcg	gcacgcgcgc	cgcctggggg	cgctctttga	tactcgccgc	5160
cgctcgagc	ccctgcggag	gtccctggag	gaggccgagg	cgagtgagg	cgaggtctgg	5220
ggctcgcttg	gcgcgctccg	ggggggcgcc	tggaggtccc	ccgagggtt	ccgcgccatg	5280
cacgagcagc	tgccgcgcct	gcaggacacc	accaacacgg	tgctggggct	gcgggcccag	5340
ccggcctacg	agcgccttcc	cgccaggtag	cagggcgtcc	tgggcgccaa	gggcgccgag	5400
cgcgccgagg	ccgtggagga	gctgggcgcg	cgctcacca	agcacaccgc	cctgtgcgcc	5460
cggtgcggg	atgaggtggt	gcgcagggtg	ccctgggaga	tgaactttga	cgccctgggg	5520
ggcctgctgg	cggagttcga	cgcgccggcc	gcggacctgg	ccccgtgggc	ggtggaggag	5580
ttccggggcg	cgcggaact	catccaatac	cgcatgggct	tatatagcgc	gtacgccagg	5640
gcccggcgcc	agacgggcgc	gggcgcggag	tcgcgcgcgc	cgccctgct	cgttgatctt	5700
cgcgcctcgc	acgccgcgc	ccgggcctct	tcttcccccg	agggccacga	ggtcgacccg	5760
cagctgctgc	gccgcgggg	cgaggcgtag	ctgcgcgcgc	gcggagaccc	ggggccccctg	5820
gtgctgcgtg	aggcgcgcag	cgctctggat	ctgcgcttcg	ccaccagctt	tttgcccccg	5880
gacggaacgc	cgctgcagta	cgcgctatgc	ttccgcgcgc	tcaccgacaa	actcggcgcg	5940
ctgctgatgc	gtcccagggc	ggcctgcgtg	cgcccccgcg	ttccgacgga	cgtcctcgaa	6000
tcggccccga	cggtcacggc	catgtacgtg	ctgaccgtcg	tgaaccggct	ccagctggcc	6060
ctcagcgacg	cccaggccgc	caactttcag	ctgttcggtc	gcttcgtgcg	ccatcgccag	6120
gcgacgtggg	gcgcctcgat	ggacgcggcg	gccgagctgt	acgtcgccct	cgtcgccacc	6180
accctcacgc	gcgagtttgg	gtgtcgctgg	gcccagctgg	gctgggcgtc	cggagcgggc	6240
gcgcgcgcgc	cgccgcgggg	cccccggggg	tcccagcgcc	actgcgtcgc	cttcaacgag	6300
aacgacgtgc	tggtcgcgct	ggtggccggc	gttccgggaac	acatctacaa	cttctggcgc	6360
ctggacctcg	ttcgccagca	cgagtacatg	cacctcacc	tcgaacgcgc	gttcgaggac	6420
gcagcggagt	ccatgctgtt	cgtccagcgc	ctgacccgcg	atcccagcgc	cgcacccgcg	6480
gtgttgccga	cgtttttgga	cggaggcccc	ccgacccggg	gcctcctggt	cggcaccgcg	6540
ctggccgact	ggcgccgggg	caagctgtcc	gaaaccgacc	cgctggcgcc	ctggcgctcg	6600
gccttgagc	tcgggaccca	gcgcggggac	gtcccggcgc	tcgggaagct	cagtccggcc	6660
caggccctgg	cgcggtgag	cgtcctcggg	cgcatgtgtc	tgccgagcgc	cgctttggcc	6720
gcgctgtgga	cctgcatgtt	tcccgacgac	tacaccgagt	acgacagctt	cgacgccttc	6780
ctggccgcac	gcctggagtc	tggccagacg	ctcgcccgcg	cgggggggcg	cgaggcgtcc	6840
ctccccgagg	ccccccacgc	cctctaccga	cccacgggcc	agcacgtggc	cgtgctggcc	6900
gcgcgcaccc	accgcacccc	cgccgcgcgc	gttacggcca	tggacctggt	tctggccgcg	6960
gtgctcctcg	gcgcgcccg	cgtggtggcg	ctccgcaaca	ccacggcctt	ctcccgcgag	7020

tcggaactgg	aactgtgcct	gacgtcttct	gactcgcgcc	ccggcggggc	ggacgcgcgc	7080
ctgcgcgacg	tctgtctcgt	cgacatcgag	acgtggggcc	tccgcctcct	ccacaccgat	7140
ctcaaccgga	tcgaaaacgc	gtgtctggcg	gcgcagctcc	cgcgcctgtc	ggcgctcctc	7200
gccgagcgcc	ctctcgccga	cgggccccc	tgcctgggtc	tctgggacat	ctccatgacc	7260
ccggtcgcgg	tcttgtggga	agccccggag	ccccccggcc	ccctgacgt	gcggtttgtg	7320
ggcagcgagg	ccaccgagga	gcttccgttt	gtggtaccg	cgggggacgt	tctcgcggcg	7380
agcgcgcgcg	acgcggaccc	cttcttcgcg	cggggccatcc	tccggcgcc	cttcgacgcc	7440
tccctcctga	cgggggagct	gttcccgga	cacccggttt	accagcgccc	cctcgcgcgac	7500
gaggcaggtc	cctctgcccc	gaccgcgcgc	cgcgaccgcg	gggaccttgc	ggggggggat	7560
ggcggatcgg	gtcccagagga	ccccgctgcc	ccccccgcac	ggcaggcgga	cccgggggtc	7620
ctcgccccc	ctctcctcac	cgacgccacc	accggcgagc	ccgtccccc	tgcgatgtgg	7680
gcctggatcc	acggcctgga	ggagctggcg	tccgacgacg	ccggcggccc	cacgcccaat	7740
ccggccccc	cccttcttcc	ccccccgcgc	accgatcagt	ccgtccccc	gtcccagtag	7800
gcaccgcggc	ccatcgggcc	ggcagctacg	gctcgcgaaa	cacgaccgag	tgtcccgcct	7860
caacaaaata	cggggcgcggt	gcccgtggcc	cctcgcgacg	acccacggcc	ctcgccaccc	7920
acaccgagtc	cccccgcgga	tgcgcgctt	cctccccgg	ccttttcgg	ctctgcgcgc	7980
gctttctccg	ccgcgctgcc	gcgcgtgcgc	aggtcccgcc	ggacgcgtgc	caaatacgcg	8040
gcgccccggg	cctcgcgcgc	ccccgagggg	tggcgtccgc	cggcaactcc	cgccccggtc	8100
gccccggctg	ccgcttcgcg	gcgtcctccc	gaccagccgc	ccaccccgga	gtccgcgcgc	8160
ccggcttggg	tttctgcttt	gccccttcgc	cccggccccg	cctcggctcg	cggggccttt	8220
cccgccccca	cgctcgcgc	catccgcct	ccgcccgcgc	aaggagcggt	cgtccccggg	8280
ggcgaccgcc	gcccgcggcg	ccgccagaca	acagcgggtc	cctcgcgcgac	acccccccgg	8340
ggccccgcgg	ccggcccaac	ccggcgccct	acgcgccccg	cggtcgcctc	gctctcggcc	8400
tctgtgaact	ccctccccc	tccccgggac	ccagccgacc	acgcggccgc	ggtttccgcc	8460
gcagcgcgcg	cagtgcgcgc	ttctccgggc	cttgccccc	cgacctcgc	cgtccagacc	8520
agcccccg	ctttggcccc	aggcccgggt	gccccgagcg	agcctttgtg	cggttgggtc	8580
gtgcccgag	cccccggttg	ccgcgcgcgc	ccacccaggt	cacccgcgac	caaaccgcgc	8640
gcgcggacgc	gcatccgcgc	gcgctcgggt	ccccagccgc	ccctgcccc	gcccgccttg	8700
ccccagccgc	ccctgcccc	gcccgccttg	ccccagccgc	ccctgcccc	gcccgccttg	8760
ccccagccgc	ccctgcccc	gcccgccttg	ccccagccgc	ccctgcccc	gcccgccttg	8820
ccccagccgc	ccctgcccgc	cgtaactcgg	acgcttacac	cccagagccg	agactcgggtc	8880
ccgacgccc	agtctccac	acatacaaac	acccacctcc	cagtgtccgc	tgtgacttcg	8940
tgggttccct	ctctcgcgct	tcacgtcgac	tccgcgcccc	cacccgcac	gctgcttcag	9000
acgtgcaca	tctcctccga	tgatgaacac	tccgacgcag	actccctgcg	tttctcggac	9060
tccgatgata	ccgaggcgct	cgacccctc	ccaccggaac	cgcacttgcc	ccccgcgcgac	9120
gaaccacccg	gcccattggc	cgccgaccac	ctccagtcac	cacactcgca	atgttgcccc	9180
ttgcccgtac	aggcaaatgc	cgtcttgcg	cgacgtacg	tacgaagcac	cggtcgtagc	9240
gccctggcgg	ttttgatccg	cgcgtgcgcg	cgcatacaac	aacagctgca	gcggaccgcg	9300
agggcgctgt	ttcagcgcag	caacgccgtg	ctgaccagcc	tacatcacgt	gcgcgatgta	9360
ctgggttag						9369

<210> 191

<211> 3591

<212> DNA

<213> HSV2

<400> 191

atggacacca	agcccaaaac	gacgaccacc	gtcaagggtc	ctccggggcc	gatgggggtac	60
gtgtatggcc	gcgcctgtcc	ggccgaaggc	ctggagctgc	tgtcgtact	gtcggcgcg	120
agcggcgacg	cggacgtcgc	cgtcgcgcgc	ctcatcgctc	gcctgaccgt	ggagagcggg	180
tttgaggcca	acgtggccgc	ggtcgtcggg	tcccgcaaga	cgggccttgg	ggggaccgcg	240
gtgtccctca	agcttatgcc	ctcgcaactac	agcccgctcc	tttacgtctt	tcacggcggc	300
cgccacctgg	ccccagcac	ccaggccccc	aacctgacgc	ggctctgcga	gcgggcgcgc	360
ccgcattttg	gcttcgcgga	ctacgcgcgc	cggccctgcg	acctaaagca	cgagacgacg	420
ggggacgcgc	tgtgcgagcg	gctcggactg	gaccgggacc	gggccttgct	gtatctggtg	480
atcacggagg	gcttccggga	ggccgtgtgc	atcagcaaca	ccttccctgca	tctgggcggg	540
atggacaagg	taaccatcgg	ggacgcggag	gttcaccgca	tccccgtgta	tccgctgcag	600
atgtttatgc	cggatttttag	ccgggtcacc	gccgatccgt	ttaactgcaa	ccaccgatcg	660

```

atcggggaga atttcaacta ccccttccg tttttcaacc gccccctcgc ccgcctgctg 720
ttcgaggcgg tcgtcggggc cgccgcgctg gccctgcgtg cccggaacgt ggacgcgctg 780
gcccgcgcgg ccgcccacct ggcggttcgac gaaaaccacg agggcgcggc cctccccgcc 840
gacattacgt tcacggcctt cgaggccagc cagggttaagc cccagcgggg cgcgcgcgac 900
gccgggaaca agggcccggc gggcggttt gagcagcgcc tggcctcggg gatggcggg 960
gacgcgccc tggcctgga gtccatcgtg tcgatggcg tctttgacga gccgcccccc 1020
gacatcacca cgtggccct gcttgagggc caggaaacgc ccgcggcccg cgccggcgcc 1080
gtcggggcgt acctggcgcg cgccgcgggc ctctggtgtt cagcaccaat 1140
tcggccctcc acctgaccga ggtggacgac gccgggcccg cggaccccaa ggaccacagc 1200
aagccctcct tttaccgctt ctttctcgtg cccggcacgc acgtggcggc caaccgcgag 1260
ctggaccgcg agggccacgt ggtgccggg tacgagggg gccccacggc gccctcgtc 1320
ggcggaaccc aggagtgcg cgcgagcac ctggccatgc tgtcggggt ctctccggc 1380
ctgctggcca agatgctgtt ttacctggag cgctgcgac gcgagtcac cgtcgggcgc 1440
caggagatgg acgtgtttcg ctacgtcgcg gactcgggcc agaccgacgt gccctgcaac 1500
ctgtgtacct tcgagaccg ccacgcctgc gcgcacacga cgtgatgcg gctcggggc 1560
cgcccccca agttcgcgag cgccggcccg ggggccattg gggctctcgg aacgatgaac 1620
agcgcgta ca gcgactgcga cgtgctggg aactacgcc ccttctcggc cctgaagcg 1680
gcgagcgggt ccgagaacac ccggaccatc atgcaggaga cgtaccgcgc ggccaccgag 1740
cgcgatcatg ccgagctcga ggccctgcag tacgtggacc aggcgggtgc caccggcctg 1800
gggaggctgg agacgattat cggcaaccgc gaggccctgc acacggtggt gaacaacatc 1860
aagcagctgg tggatcgcg ggtggaacag ctgatgcgca acctgatcga gggcgcaac 1920
tttaagttcc gcgacggtct ggccgaggcc aaccacgcca tgtcgtgtc gctggaccgc 1980
tacacctgcg ggccgtgcc cctgctgcag ctctcgcgc ggcggtccaa cctcgccgtg 2040
taccaggacc tggccctgag ccagtgcac ggggtgttcg ccgggcagtc ggtcgagggg 2100
cgcaactttc gcaatcaatt ccaaccggtg ctgcggcggc gcgtcatgga cctgttcaac 2160
aacgggttcc tgtcggcgaa gacgtgcgac gtcgcgtctc cggagggggc ggctatctgc 2220
gccccagcc taacggccgg ccagacggcc cccgcgaga gcagcttcga gggcgacgtt 2280
gcccgcgtga cccttgggtt tcccaaggaa ctgcgcgtca agagccgcgt gttgttcgcg 2340
ggcgcgagcg ccaacgcgtc cgaggccgcc aaggcgcggg tcgccagcct ccagagcgcc 2400
taccagaagc ccgacaagcg cgtggacatc ctctcggac cgctgggctt cctgctgaag 2460
cagttccacg cggtcatctt tcccaacggc aagcccccg ggtcgaacca gccgaaccgc 2520
cagtgttct ggaccgcct ccagcgcaac cagctccccg ccggctctt gtcgcgcgag 2580
gacatagaga catcgcggt cattaagagg ttctccctgg actacggcgc gattaacttt 2640
atcaacctgg ccccaataa cgtgagcgaa ctggccatgt actacatggc caaccagatc 2700
ttgaggtact gcgaccatc gacgtacttc ataaacacc tcacggcgt catcgccggg 2760
tcccgcgctc ccccgagct gcaggcgggc gccgcctggg ccccgaggg cggcgcgggc 2820
cttgaggccg ggcccgcgc gctgatggac tccctcgac cgacccccg cgcctggacg 2880
tccatgttcg cgagctgcaa cctgctgcg cccgtgatgg ccgcgcgcc gatggtcgtg 2940
ctgggcctca gcatcagcaa atactacggc atggccggca acgaccgcgt gtttcaggcc 3000
gggaactggg ccagcctgct gggcggaaca aacgcgtgcc cgctgctgat ctttgaccgc 3060
acccgcaagt ttgtcctggc ctgcccagg gccgggtttg tgtgcgcggc gtcgagctc 3120
ggggcgggg cccacgagca ttcgctgtgc gagcagctcc ggggcatcat tgcggagggc 3180
ggggcgggc tcgccagcag cgtgtttgtg gcgaccgtca agagcctgg ccccgacc 3240
cagcagctgc agatcgagga ctggctggcg ctgctggagg acgagtacct gagcaggaa 3300
atgatggagt ttaccacccg cgccctggag cgcgggcatg gcgagtgtc gaccgacgcg 3360
gccctggagg tggccacga ggcgaggcc ctggtcagcc agctcggcgc cgccggggag 3420
gtgtttaact ttggggattt cggcgacgag gatgaccacg cggcgtcgtt cggcggtttg 3480
gcgcgcgcg cgggggcggc cggcgtcgcc cgaagcggg cgtttcacgg cgacgatccg 3540
tttggggagg gaccaccgga gaaaaaggac ctgacgttgg atatgctcta g 3591

```

<210> 192
 <211> 1758
 <212> DNA
 <213> HSV2

<400> 192
 atggaccggt actacccttt cgacgcgctg gacgtttggg aacacaggcg cttcatcgtc 60
 gccgactcca ggagcttcat caccgccgag ttcccccggg acttctggat gttgcccggt 120


```

ttcaacatcc cccgggagac ggccggcgag cgggcggcag tgctgcaggc ccagcgcacc 180
gcggccgcgg cggccctgga gaacgcgcgc ctccaggccg ccgagctgcc cgtcgacatc 240
gagcgccgga tacgcccgat cgagcagcag gtgcatcaca tcgcccagcg cctggaggcg 300
ctggagaccg cggcggccgc ggccgaagag gcgcatgccg cgcgggacgc cgaggcgagg 360
ggggagggcg ctgcggacgg ggcagcgccg tcgcccaccg cgggccccgc cgccgcggag 420
atggagggtc agatcgtacg caacgacccg ccgctacgat acgataccaa cctccccgtg 480
gatctgctac acatggtgta cgcgggcccgc ggggcccgcg gttcgtcggg agtcgtcttt 540
ggtacctggt accgcacgat ccaggaacgc accatcgcgg acttccccct gaccaccgcg 600
agcgccgact ttcgagacgg gcgcatgtcc aagaccttca tgaccgcgct ggtcctgtct 660
ctgcagtcgt gcggccggct gtacgtgggc cagcgccact attccgcctt cgagtgcgcc 720
gtgctgtgtc tgtatctgct gtaccgaacc acccacgagt cctccccoga tcgcgatcgc 780
gctcccgttg cgttcgggga cctgctggcc cgcctgccgc gctacctggc gcgtctggcc 840
gcggtaatcg gcgacgagag cggacgcccg cagtaccgct accgcgacga caagctgccc 900
aaagcgcagt tcgcgcggc cggcggccgc tacgagcacg gggccctggc caccacgctc 960
gtgatcgcca cgttggtgag ccacgggggtg ctaccggcgg ccccgggcga cgttccccga 1020
gacaccagca cccgctgaa ccccgacgac gtggcccacc gcgacgacgt caaccgcgc 1080
gccgccgctg ttttggcacg cggccacaac ctcttctgtg gggaggacca gacgtgctg 1140
cgggcgaccg ccaacaccat tacggccctg gccgtgcttc ggcggtcctc cgcgacggc 1200
aacgtgtacg cggaccgcct cgacaaccgc ctgcagctgg gcatgctgat cccgggagcc 1260
gtcccggcgg aggccatcgc tcggggggcg tcgggattgg actcgggcgc cataaaaagc 1320
ggcgacaaca acctggaggc gctgtgcgtt aactatgtac ttccgctgta tcaggcagac 1380
cccacggtcg agctgaccca gttgtttccg gggctggccg ccctgtgcct ggacgcccag 1440
gcggggcggc cactggcgtc gacgaggcgc gtggtggata tgcgtcggg cgcccgccag 1500
gcggcgctcg tgcgcctcac cgcgtggag ctcacaaacc gcacccgcac aaacaccacc 1560
cctgtggggg agattattaa cgccacgat gccttgggga tacaatacga acaggggcct 1620
gggctgctcg ccagcagggc acgcatcggc ttggcgtaaa acaccaagcg attcgccacg 1680
ttcaacgtgg cgacgcgacta cgacctgttg tacttttgtg gtctcggggt cattccccag 1740
tacctgtccg tggcctag 1758

```

<210> 193

<211> 3957

<212> DNA

<213> HSV2

<400> 193

```

atgtcggcgg agcagcgga gaagaagaag acgacgacga cgacgcaggg ccgcgggggc 60
gaggtcgcga tggcgagcga ggacggggga cgtctccggg ccgcggcgga gacgaccggc 120
ggccccgat ctccgatcc agccgacgga cgcgcgcca cccgaacc ccgaccgtcg 180
ccgcgcgcg gcccggtt cgggtggcac ggtggcgccg aggagaacga agacgaggcc 240
gacgacgccc ccgcgatgc cgatgccgac gaggcggccc cggcgctcgg ggaggccgtc 300
gacgagcctg ccgcggacgg cgtcgtctcg ccgcggcagc tggccctgct ggctcgtatg 360
gtggacgagg ccgttcgcac gatcccgctg cccccccgg agcgcgacgg cgcgcaagaa 420
gaagcgcccc gctcgccctt tcgcgcggg accccctcca tgcgcgccga ttatggcgag 480
gagaacgacg acgacgacga cgacgacgat gacgacgacc gcgacgcggg ccgctgggtc 540
cgcgacccgg agacgacgtc cgcggtccgc gggcggtacc cggaccccat ggccagcctg 600
tcgcgcgac ccccgcgccc ccgcgacac caccaccacc accaccaccg ccgcccggcg 660
gcccccgcc ggcgctcggc cgcctctgac tcatcaaaat ccggatcctc gtcgtcggcg 720
tctccgctc cctcctcgc ctcctcctc tcgtctgcat ccgctcctc gtctgacgac 780
gacgacgacg acgacgcgc ccgcgcccc cagaccacgc cgcgggcggg 840
accctcgggc cggacgacga ggaggcggg gtgcccgcga gggccccggg ggccggcgccc 900
cgcccgagcc cgccagggg cgagcccgc ccggcccga ccccgcggc gaccgcgggc 960
cgctggagc gccgcgggc ccgcgcggc gtggccggcc gcgacgccac gggccgcttc 1020
acggccgggc ggccccggcg ggtcgagctg gacgccgacg cggcctccgg cgccttctac 1080
gcgcgtacc gcgacgggta cgtcagcggg gagccgtggc ccggggccgg cccccgccc 1140
ccggggcgcg tgctgtacgg cgggtgggc gacagccgcc ccggcctctg gggggcgccc 1200
gaggcgagg aggcgcgggc ccggttcgag gcctcggggc cccggcgccc cgtgtggggc 1260
cccagctgg gcgacgcggc gcagcagtag gccctgatca cgcggctgct gtacacgccg 1320
gacgcggagg cgatggggtg gctccagaac ccgcgcgtgg cggccgggga cgtggcgctg 1380

```



```

gaccaggcct gcttccggat ctcgggcgcg gcgcgcaaca gcagctcctt catctccggc 1440
agcgtggcgc gggccgtgcc ccacctgggg tacgccatgg cggcggggcg cttcgggtgg 1500
ggcctggcgc acgtggcggc cgccgtggcc atgagccgcc gctacgaccg cgcgcagaag 1560
ggcttcctgc tgaccagcct gcgcgcgcc tacgcgcccc tgcctggcgc cgagaacgcg 1620
gcgctgaccg gggcgcgaaac ccccgacgac ggcggcgacg ccaaccgcca cgacggcgac 1680
gacgcccgcg ggaagcccgc cgccgcgcgc gccccgttgc cgtcggcggc ggcgtcgccg 1740
gccgacgagc gcgcggtgcc cgccggctac ggcgcgcgcg ggggtgctgc cgccctgggg 1800
cgctgagcgc ccgcgcgcgc ctccgcgcgc gccggggccg acgacgacga cgacgacgac 1860
ggcgccggcg gtggtggcgc cggccggcgc gcggaggcgc gccgcgtggc cgtggagtgc 1920
ctggccgcct gccgcgggat cctggaggcg ctggcgagg gcttcgacgc cgacctggcg 1980
gccgtgcgcg ggctggcgcg agcccggccc gccgcgcccc cgcccccggg gcccgcgggc 2040
ggggccgcgc cgccgcacgc cgacgcgcgc cgctgcgcgc cctgggtgcg cgagctgcgc 2100
ttcgtgcgcg acgcgtgggt gctgatgcgc ctgcgcgggg acctgcgcgt ggccggcggc 2160
agcgaggcgc ccgtggcggc cgtgcgcgcc gtgagcctgg tcgcgggggc cctgggcccg 2220
gcgctgcgcg ggagcccgcg cctgctgagc tccgcgcgcg ccgcgcgcgc ggacctgctc 2280
ttccagaacc agagcctgcg cccctgctgc gcgcacaccg tcgcgcgggc cgactcgctc 2340
gccgcgcgcg cctccgcgcg gcgggaggcg cgcaagcgca agagcccgcc cccggccagg 2400
ggcgccgcgc ggcgcgcgcc ggcgcgcgcg aagaagagcc gcgcggagcg ccccgcccc 2460
ggggccgcgc ctcccgcggg ggcgcgcgcc ccgcgcgcgc cgacgcgcgc gcccgggcg 2520
ccgcgcgcgc cggcgtgac ccgcgcgcgc gccgagggcc ccgacccgca gggcggtgg 2580
cgccgccagc cgccggggcc cagccacacg ccggcgccct cgccgcgcgc cctggaggcc 2640
tactgcgcgc cgccggcggt ggccgagctc acggaccacc cgctcttccc cgcgccgtgg 2700
cgcccgcccc tcatgttcga cccgcgcgcg ctggcctcgc tggccgcgcg ctgcgcgcgc 2760
ccgccccccg gcgcgcgcgc cgccgccttc ggcccgctgc gcgcctcggg cccgctgcgc 2820
cgcgcgcgcg cctggatgcg ccagggtgcc gacccgagg acgtgcgcgt ggtgatctc 2880
tactgcgcgc tgccggcgca ggacctggcc gcgggcgcgc ccggggggcg gcccccccg 2940
gagtggtccg ccagcgcgcg cgggctgtcc tgcctgctgg cgccctggg caaccggctc 3000
tgccggcccc ccacggccgc ctggggcggg aactggaccg gcgccccga cgtctcggcg 3060
ctgggcgcgc agggcggtgt gctgctgtcc acgcgggacc tggccttcgc cggcgccgtg 3120
gagttcctgg ggctgctggc cggcgctgc gaccgcgcgc tcatcgctgt caacgccgtg 3180
cgcgcccgcg actggccgcg cgacgggccc gtggtctcgc ggcagcacgc ctacctggcc 3240
tgcgaggtgc tgcccgccgt gcagtgcgcg gtgcgctggc cggcggcgcg ggacctgcgc 3300
cgcacccgtc tggcctcccg ccgcgtgttc gggccggggg tcttcgcgcg cgtggaggcc 3360
gcgcacgcgc gcctgtaccc cgacgcgcgc cgctgcgcgc tctgcgcgcg ggccaacgtg 3420
cggataccgc tgccgcacgc cttcggcccc gacacgctgg tgcccatgtc cccgcgcgag 3480
taccgcgcgc ccgtgctccc ggcgtgggac ggcggggccg ccgcctcggg cgcgggcgac 3540
gccatggcgc ccggcgcgcc ggacttctgc gaggacgagg cgcactcgca ccgcgcctgc 3600
gcgcgtggg gccctggcgc gcgcgtgcgc ccgcttacg tggcgctggg gcgcgacgc 3660
gtgcgcggcg ccccgcgga gctgcgcgg cgcgggcggg agttctgcgc gcgggcgtg 3720
ctcgagcccg acggcgacgc gccccgctg gtgctgcgcg acgacgcgga cgcgggccc 3780
ccccgcgca tacgtgggc gtcggccgcg ggccgcgcgc ggacggtgct ggccgcggcg 3840
ggcgggcgcg tggaggtggg ggggaccgcc gcggggctgg ccacgcgcgc gaggcgcgag 3900
cccgtggaca tggacgcgga gctggaggac gacgacgacg gactgtttgg ggagtga 3957

```

<210> 194

<211> 1242

<212> DNA

<213> HSV2

<400> 194

```

atggcagaca tccccccgga cccgcccgcg ctcaaacaga cgctgcgaa tcatgctccc 60
ccatccccac ccccggttcc acggaagcgc agacgcccgc tcctccccag ctgctcgga 120
tctgagggta agcccgacac agaatcgga tcctcctcga ccgagtcgtc cgaggatgag 180
gcgggagacc tacgcggcgc gcgcgcgcgc tcccccggg agctcggggg gaggtatatt 240
ttgatctgt cggcagaatc gaccacgggg acggaatcgc agggaacggg gccgtcggac 300
gacgatgatg atgatgcgtc agacggctgg ttggttgaca ccccccccg caaatccaag 360
cgaccccgaa tcaacctgcg attaacgagc tcccccgacc ggcgtgcggg tgtggtttt 420
cccgaggtgt ggagaagcga cagacctatc cgcgcggcgc aaccccgagg cccggccagt 480

```

```

cttccgggga tcgcgcacgc gcaccggcgc tctgctcgcc aggcccagat gcggagcgga 540
gccgcctgga cgcttgatct gcattacata cgccagtgcg tcaaccagct ctttcggatc 600
ctgcgtgccg ccccgaaacc gcccggcagc gccaaaccgc tcgcccacct ggtgcgagac 660
tgctacctca tgggctactg ccggaccgcg ctggggcgcg gcacgtgggg ccgcctgctg 720
cagatctcgg gcggaacctg ggacgtgcgc ctgcgaaacg caatccggga ggtcgaggcg 780
cattttgaac ccgccgccga gcccggtgtgc gagctgccct gtctgaacgc caggcgttac 840
ggccccgagt gtgatgttg caatctcgag accaacggcg gctcgacgag cgatgatgag 900
atatcggatg cgacggactc ggacgatacc ctccggtccc attccgacac ggaggggggg 960
ccctccccgg ccggccggga gaaccggaa tccgcgtccg gcggggctat cgcggctcgg 1020
ctggagtgtg agtttgggac gtttgactgg acgtccgagg agggctccca gccctggctg 1080
tccgcggttg tcgccgatac cagctccgcc gaacgctctg gcctacccgc cccgggcgcg 1140
tgtcgcgcaa cggaagcccc agaacgcgag gacgggtgcc gaaaaatgcg cttccccgcc 1200
gcctgccccct atccctgcgg ccacacattt ctccggccat ga 1242

```

<210> 195

<211> 318

<212> PRT

<213> HSV2

<400> 195

```

Met Ile Thr Asp Cys Phe Glu Ala Asp Ile Ala Ile Pro Ser Gly Ile
          5              10              15

```

```

Ser Arg Pro Asp Ala Ala Ala Leu Gln Arg Cys Glu Gly Arg Val Val
          20              25              30

```

```

Phe Leu Pro Thr Ile Arg Arg Gln Leu Ala Leu Ala Asp Val Ala His
          35              40              45

```

```

Glu Ser Phe Val Ser Gly Gly Val Ser Pro Asp Thr Leu Gly Leu Leu
          50              55              60

```

```

Leu Ala Tyr Arg Arg Arg Phe Pro Ala Val Ile Thr Arg Val Leu Pro
          65              70              75              80

```

```

Thr Arg Ile Val Ala Cys Pro Val Asp Leu Gly Leu Thr His Ala Gly
          85              90              95

```

```

Thr Val Asn Leu Arg Asn Thr Ser Pro Val Asp Leu Cys Asn Gly Asp
          100              105              110

```

```

Pro Val Ser Leu Val Pro Pro Val Phe Glu Gly Gln Ala Thr Asp Val
          115              120              125

```

```

Arg Leu Glu Ser Leu Asp Leu Thr Leu Arg Phe Pro Val Pro Leu Pro
          130              135              140

```

```

Thr Pro Leu Ala Arg Glu Ile Val Ala Arg Leu Val Ala Arg Gly Ile
          145              150              155              160

```

```

Arg Asp Leu Asn Pro Asp Pro Arg Thr Pro Gly Glu Leu Pro Asp Leu
          165              170              175

```

```

Asn Val Leu Tyr Tyr Asn Gly Ala Arg Leu Ser Leu Val Ala Asp Val
          180              185              190

```

```

Gln Gln Leu Ala Ser Val Asn Thr Glu Leu Arg Ser Leu Val Leu Asn
          195              200              205

```

Met Val Tyr Ser Ile Thr Glu Gly Thr Thr Leu Ile Leu Thr Leu Ile
 210 215 220

Pro Arg Leu Leu Ala Leu Ser Ala Gln Asp Gly Tyr Val Asn Ala Leu
 225 230 235 240

Leu Gln Met Gln Ser Val Thr Arg Glu Ala Ala Gln Leu Ile His Pro
 245 250 255

Glu Ala Pro Met Leu Met Gln Asp Gly Glu Arg Arg Leu Pro Leu Tyr
 260 265 270

Glu Ala Leu Val Ala Trp Leu Ala His Ala Gly Gln Leu Gly Asp Ile
 275 280 285

Leu Ala Leu Ala Pro Ala Val Arg Val Cys Thr Phe Asp Gly Ala Ala
 290 295 300

Val Val Gln Ser Gly Asp Met Ala Pro Val Ile Arg Tyr Pro
 305 310 315

<210> 196
 <211> 413
 <212> PRT
 <213> HSV2

<400> 196

Met Ala Asp Ile Pro Pro Asp Pro Pro Ala Leu Asn Thr Thr Pro Ala
 5 10 15

Asn His Ala Pro Pro Ser Pro Pro Pro Gly Ser Arg Lys Arg Arg Arg
 20 25 30

Pro Val Leu Pro Ser Ser Ser Glu Ser Glu Gly Lys Pro Asp Thr Glu
 35 40 45

Ser Glu Ser Ser Ser Thr Glu Ser Ser Glu Asp Glu Ala Gly Asp Leu
 50 55 60

Arg Gly Gly Arg Arg Arg Ser Pro Arg Glu Leu Gly Gly Arg Tyr Phe
 65 70 75 80

Leu Asp Leu Ser Ala Glu Ser Thr Thr Gly Thr Glu Ser Glu Gly Thr
 85 90 95

Gly Pro Ser Asp Asp Asp Asp Asp Ala Ser Asp Gly Trp Leu Val
 100 105 110

Asp Thr Pro Pro Arg Lys Ser Lys Arg Pro Arg Ile Asn Leu Arg Leu
 115 120 125

Thr Ser Ser Pro Asp Arg Arg Ala Gly Val Val Phe Pro Glu Val Trp
 130 135 140

Arg Ser Asp Arg Pro Ile Arg Ala Ala Gln Pro Gln Ala Pro Ala Ser
 145 150 155 160

Leu Pro Gly Ile Ala His Ala His Arg Arg Ser Ala Arg Gln Ala Gln
 165 170 175
 Met Arg Ser Gly Ala Ala Trp Thr Leu Asp Leu His Tyr Ile Arg Gln
 180 185 190
 Cys Val Asn Gln Leu Phe Arg Ile Leu Arg Ala Ala Pro Asn Pro Pro
 195 200 205
 Gly Ser Ala Asn Arg Leu Arg His Leu Val Arg Asp Cys Tyr Leu Met
 210 215 220
 Gly Tyr Cys Arg Thr Arg Leu Gly Pro Arg Thr Trp Gly Arg Leu Leu
 225 230 235 240
 Gln Ile Ser Gly Gly Thr Trp Asp Val Arg Leu Arg Asn Ala Ile Arg
 245 250 255
 Glu Val Glu Ala His Phe Glu Pro Ala Ala Glu Pro Val Cys Glu Leu
 260 265 270
 Pro Cys Leu Asn Ala Arg Arg Tyr Gly Pro Glu Cys Asp Val Gly Asn
 275 280 285
 Leu Glu Thr Asn Gly Gly Ser Thr Ser Asp Asp Glu Ile Ser Asp Ala
 290 295 300
 Thr Asp Ser Asp Asp Thr Leu Ala Ser His Ser Asp Thr Glu Gly Gly
 305 310 315 320
 Pro Ser Pro Ala Gly Arg Glu Asn Pro Glu Ser Ala Ser Gly Gly Ala
 325 330 335
 Ile Ala Ala Arg Leu Glu Cys Glu Phe Gly Thr Phe Asp Trp Thr Ser
 340 345 350
 Glu Glu Gly Ser Gln Pro Trp Leu Ser Ala Val Val Ala Asp Thr Ser
 355 360 365
 Ser Ala Glu Arg Ser Gly Leu Pro Ala Pro Gly Ala Cys Arg Ala Thr
 370 375 380
 Glu Ala Pro Glu Arg Glu Asp Gly Cys Arg Lys Met Arg Phe Pro Ala
 385 390 395 400
 Ala Cys Pro Tyr Pro Cys Gly His Thr Phe Leu Arg Pro
 405 410

<210> 197
 <211> 1318
 <212> PRT
 <213> HSV2

<400> 197
 Met Ser Ala Glu Gln Arg Lys Lys Lys Lys Thr Thr Thr Thr Thr Gln
 5 10 15

Gly Arg Gly Ala Glu Val Ala Met Ala Asp Glu Asp Gly Gly Arg Leu
 20 25 30
 Arg Ala Ala Ala Glu Thr Thr Gly Gly Pro Gly Ser Pro Asp Pro Ala
 35 40 45
 Asp Gly Pro Pro Pro Thr Pro Asn Pro Asp Arg Arg Pro Ala Ala Arg
 50 55 60
 Pro Gly Phe Gly Trp His Gly Gly Pro Glu Glu Asn Glu Asp Glu Ala
 65 70 75 80
 Asp Asp Ala Ala Ala Asp Ala Asp Ala Asp Glu Ala Ala Pro Ala Ser
 85 90 95
 Gly Glu Ala Val Asp Glu Pro Ala Ala Asp Gly Val Val Ser Pro Arg
 100 105 110
 Gln Leu Ala Leu Leu Ala Ser Met Val Asp Glu Ala Val Arg Thr Ile
 115 120 125
 Pro Ser Pro Pro Pro Glu Arg Asp Gly Ala Gln Glu Glu Ala Ala Arg
 130 135 140
 Ser Pro Ser Pro Pro Arg Thr Pro Ser Met Arg Ala Asp Tyr Gly Glu
 145 150 155 160
 Glu Asn Asp Asp Asp Asp Asp Asp Asp Asp Asp Asp Arg Asp Ala
 165 170 175
 Gly Arg Trp Val Arg Gly Pro Glu Thr Thr Ser Ala Val Arg Gly Ala
 180 185 190
 Tyr Pro Asp Pro Met Ala Ser Leu Ser Pro Arg Pro Pro Ala Pro Arg
 195 200 205
 Arg His His His His His His His Arg Arg Arg Arg Ala Pro Arg Arg
 210 215 220
 Arg Ser Ala Ala Ser Asp Ser Ser Lys Ser Gly Ser Ser Ser Ser Ala
 225 230 235 240
 Ser Ser Ala Ser Ser Ser Ala Ser Ser Ser Ser Ser Ala Ser Ala Ser
 245 250 255
 Ser Ser Asp Asp Asp Asp Asp Asp Asp Ala Ala Arg Ala Pro Ala Ser
 260 265 270
 Ala Ala Asp His Ala Ala Gly Gly Thr Leu Gly Ala Asp Asp Glu Glu
 275 280 285
 Ala Gly Val Pro Ala Arg Ala Pro Gly Ala Ala Pro Arg Pro Ser Pro
 290 295 300
 Pro Arg Ala Glu Pro Ala Pro Ala Arg Thr Pro Ala Ala Thr Ala Gly
 305 310 315 320

Arg Leu Glu Arg Arg Arg Ala Arg Ala Ala Val Ala Gly Arg Asp Ala
 325 330 335
 Thr Gly Arg Phe Thr Ala Gly Arg Pro Arg Arg Val Glu Leu Asp Ala
 340 345 350
 Asp Ala Ala Ser Gly Ala Phe Tyr Ala Arg Tyr Arg Asp Gly Tyr Val
 355 360 365
 Ser Gly Glu Pro Trp Pro Gly Ala Gly Pro Pro Pro Pro Gly Arg Val
 370 375 380
 Leu Tyr Gly Gly Leu Gly Asp Ser Arg Pro Gly Leu Trp Gly Ala Pro
 385 390 395 400
 Glu Ala Glu Glu Ala Arg Ala Arg Phe Glu Ala Ser Gly Ala Pro Ala
 405 410 415
 Pro Val Trp Ala Pro Glu Leu Gly Asp Ala Ala Gln Gln Tyr Ala Leu
 420 425 430
 Ile Thr Arg Leu Leu Tyr Thr Pro Asp Ala Glu Ala Met Gly Trp Leu
 435 440 445
 Gln Asn Pro Arg Val Ala Pro Gly Asp Val Ala Leu Asp Gln Ala Cys
 450 455 460
 Phe Arg Ile Ser Gly Ala Ala Arg Asn Ser Ser Ser Phe Ile Ser Gly
 465 470 475 480
 Ser Val Ala Arg Ala Val Pro His Leu Gly Tyr Ala Met Ala Ala Gly
 485 490 495
 Arg Phe Gly Trp Gly Leu Ala His Val Ala Ala Ala Val Ala Met Ser
 500 505 510
 Arg Arg Tyr Asp Arg Ala Gln Lys Gly Phe Leu Leu Thr Ser Leu Arg
 515 520 525
 Arg Ala Tyr Ala Pro Leu Leu Ala Arg Glu Asn Ala Ala Leu Thr Gly
 530 535 540
 Ala Arg Thr Pro Asp Asp Gly Gly Asp Ala Asn Arg His Asp Gly Asp
 545 550 555 560
 Asp Ala Arg Gly Lys Pro Ala Ala Ala Ala Ala Pro Leu Pro Ser Ala
 565 570 575
 Ala Ala Ser Pro Ala Asp Glu Arg Ala Val Pro Ala Gly Tyr Gly Ala
 580 585 590
 Ala Gly Val Leu Ala Ala Leu Gly Arg Leu Ser Ala Ala Pro Ala Ser
 595 600 605
 Ala Pro Ala Gly Ala Asp Asp Asp Asp Asp Asp Asp Gly Ala Gly Gly
 610 615 620
 Gly Gly Gly Gly Arg Arg Ala Glu Ala Gly Arg Val Ala Val Glu Cys

625		630		635		640
Leu Ala Ala Cys Arg Gly Ile Leu Glu Ala Leu Ala Glu Gly Phe Asp						
	645			650		655
Gly Asp Leu Ala Ala Val Pro Gly Leu Ala Gly Ala Arg Pro Ala Ala						
	660			665		670
Pro Pro Arg Pro Gly Pro Ala Gly Ala Ala Ala Pro Pro His Ala Asp						
	675			680		685
Ala Pro Arg Leu Arg Ala Trp Leu Arg Glu Leu Arg Phe Val Arg Asp						
	690			695		700
Ala Leu Val Leu Met Arg Leu Arg Gly Asp Leu Arg Val Ala Gly Gly						
	705			710		715
Ser Glu Ala Ala Val Ala Ala Val Arg Ala Val Ser Leu Val Ala Gly						
	725			730		735
Ala Leu Gly Pro Ala Leu Pro Arg Ser Pro Arg Leu Leu Ser Ser Ala						
	740			745		750
Ala Ala Ala Ala Ala Asp Leu Leu Phe Gln Asn Gln Ser Leu Arg Pro						
	755			760		765
Leu Leu Ala Asp Thr Val Ala Ala Ala Asp Ser Leu Ala Ala Pro Ala						
	770			775		780
Ser Ala Pro Arg Glu Ala Arg Lys Arg Lys Ser Pro Ala Pro Ala Arg						
	785			790		795
Ala Pro Pro Gly Gly Ala Pro Arg Pro Pro Lys Lys Ser Arg Ala Asp						
	805			810		815
Ala Pro Arg Pro Ala Ala Ala Pro Pro Ala Gly Ala Ala Pro Pro Ala						
	820			825		830
Pro Pro Thr Pro Pro Pro Arg Pro Pro Arg Pro Ala Ala Leu Thr Arg						
	835			840		845
Arg Pro Ala Glu Gly Pro Asp Pro Gln Gly Gly Trp Arg Arg Gln Pro						
	850			855		860
Pro Gly Pro Ser His Thr Pro Ala Pro Ser Ala Ala Ala Leu Glu Ala						
	865			870		875
Tyr Cys Ala Pro Arg Ala Val Ala Glu Leu Thr Asp His Pro Leu Phe						
	885			890		895
Pro Ala Pro Trp Arg Pro Ala Leu Met Phe Asp Pro Arg Ala Leu Ala						
	900			905		910
Ser Leu Ala Ala Arg Cys Ala Ala Pro Pro Pro Gly Gly Ala Pro Ala						
	915			920		925
Ala Phe Gly Pro Leu Arg Ala Ser Gly Pro Leu Arg Arg Ala Ala Ala						
	930			935		940

Trp Met Arg Gln Val Pro Asp Pro Glu Asp Val Arg Val Val Ile Leu
 945 950 955 960
 Tyr Ser Pro Leu Pro Gly Glu Asp Leu Ala Ala Gly Arg Ala Gly Gly
 965 970 975
 Gly Pro Pro Pro Glu Trp Ser Ala Glu Arg Gly Gly Leu Ser Cys Leu
 980 985 990
 Leu Ala Ala Leu Gly Asn Arg Leu Cys Gly Pro Ala Thr Ala Ala Trp
 995 1000 1005
 Ala Gly Asn Trp Thr Gly Ala Pro Asp Val Ser Ala Leu Gly Ala Gln
 1010 1015 1020
 Gly Val Leu Leu Leu Ser Thr Arg Asp Leu Ala Phe Ala Gly Ala Val
 1025 1030 1035 1040
 Glu Phe Leu Gly Leu Leu Ala Gly Ala Cys Asp Arg Arg Leu Ile Val
 1045 1050 1055
 Val Asn Ala Val Arg Ala Ala Asp Trp Pro Ala Asp Gly Pro Val Val
 1060 1065 1070
 Ser Arg Gln His Ala Tyr Leu Ala Cys Glu Val Leu Pro Ala Val Gln
 1075 1080 1085
 Cys Ala Val Arg Trp Pro Ala Ala Arg Asp Leu Arg Arg Thr Val Leu
 1090 1095 1100
 Ala Ser Gly Arg Val Phe Gly Pro Gly Val Phe Ala Arg Val Glu Ala
 1105 1110 1115 1120
 Ala His Ala Arg Leu Tyr Pro Asp Ala Pro Pro Leu Arg Leu Cys Arg
 1125 1130 1135
 Gly Ala Asn Val Arg Tyr Arg Val Arg Thr Arg Phe Gly Pro Asp Thr
 1140 1145 1150
 Leu Val Pro Met Ser Pro Arg Glu Tyr Arg Arg Ala Val Leu Pro Ala
 1155 1160 1165
 Leu Asp Gly Arg Ala Ala Ala Ser Gly Ala Gly Asp Ala Met Ala Pro
 1170 1175 1180
 Gly Ala Pro Asp Phe Cys Glu Asp Glu Ala His Ser His Arg Ala Cys
 1185 1190 1195 1200
 Ala Arg Trp Gly Leu Gly Ala Pro Leu Arg Pro Val Tyr Val Ala Leu
 1205 1210 1215
 Gly Arg Asp Ala Val Arg Gly Gly Pro Ala Glu Leu Arg Gly Pro Arg
 1220 1225 1230
 Arg Glu Phe Cys Ala Arg Ala Leu Leu Glu Pro Asp Gly Asp Ala Pro
 1235 1240 1245

Pro Leu Val Leu Arg Asp Asp Ala Asp Ala Gly Pro Pro Pro Gln Ile
 1250 1255 1260

Arg Trp Ala Ser Ala Ala Gly Arg Ala Gly Thr Val Leu Ala Ala Ala
 1265 1270 1275 1280

Gly Gly Gly Val Glu Val Val Gly Thr Ala Ala Gly Leu Ala Thr Pro
 1285 1290 1295

Pro Arg Arg Glu Pro Val Asp Met Asp Ala Glu Leu Glu Asp Asp Asp
 1300 1305 1310

Asp Gly Leu Phe Gly Glu
 1315

<210> 198

<211> 419

<212> PRT

<213> HSV2

<400> 198

Met Pro Arg Val Ser Ser Ser Cys Ser Ser Ser Phe Leu Pro Ala Arg
 5 10 15

Val Arg Arg Gly Gly Arg Gly Gly Gln Val Cys Gly Arg Gly Gly Arg
 20 25 30

Arg Gly Gly Gly Gly His Gly Arg Lys Cys Gly Cys Met Pro Arg Val
 35 40 45

Phe Ser Ser Pro Ser Ser Phe Leu Pro Pro Val Pro Gly Gly Arg Gly
 50 55 60

Arg Ala Cys Val Val Ile Gln Pro Pro Ser Pro Pro Pro His Phe Pro
 65 70 75 80

Pro Ser Leu Ser Lys Phe Pro Gly Pro Trp Leu Arg Ala Gly Gly Ala
 85 90 95

Ala Asp Pro Pro Pro Pro Pro Ser Pro Ser Gln Ala Pro Ser His Ser
 100 105 110

Cys Pro Pro Pro Pro Gly Ser Gly Arg Pro Asp Val Arg Ala Leu His
 115 120 125

Asp Arg Ala Pro Leu Pro Val Asn Thr Asp Thr Leu Phe Phe Tyr Pro
 130 135 140

Pro Ala Ser Pro Pro Thr His Gln Asp Arg Glu Pro Glu Arg Gly Arg
 145 150 155 160

Ala Pro Ala Leu Phe Tyr Asp Lys Asp Gln Gln Ala Ser Gly Val Gly
 165 170 175

Ala Ala Ser Arg Ala Arg Pro Pro Ser Ser Ser Leu Pro Pro His
 180 185 190

Pro Arg Pro Pro Cys Ala Gly Glu Leu His Gln Arg Pro Thr Thr Lys
 195 200 205
 Cys Val Lys Ser Ile Thr Lys Leu Tyr Cys Lys Ile Phe Ile Asn Ile
 210 215 220
 Lys Phe Phe Phe Ser Ser Ser Phe Gln Gln Gly Gln Lys Val His Asn
 225 230 235 240
 Lys Met Leu Val Cys Val Ala Val Arg Gly Arg Val Arg Pro Pro Pro
 245 250 255
 Pro Leu Pro Pro Pro Leu Pro Val Ser Ser Pro Ser Phe Pro Pro Pro
 260 265 270
 Thr Ser Pro Cys Pro Arg Gly Ala Ser Ala Gly Gly Pro Val Gly Gly
 275 280 285
 Gly Phe Leu Arg Ala Ala Ser Arg Val Leu Ala Pro Pro Thr Pro Arg
 290 295 300
 Gly Pro Arg Gly Arg Arg Arg Pro Ala Arg Ala Arg Pro Ala Pro Glu
 305 310 315 320
 Thr Thr Gly Gly Ala Thr Gly Gly Arg Gly Ser Pro Ala Arg Pro Pro
 325 330 335
 Gly Cys Pro Gly Gln Arg Ala Cys Cys Pro Pro Gly Ser Ser Arg Gly
 340 345 350
 Cys Ser Thr Pro Thr Arg Gly Ala Ser Gly Ser Arg Gly Ala Ser Arg
 355 360 365
 Ser Cys Ala Arg Ala Arg Gly Ala Gly Cys Gly Gly Ser Thr Arg Gly
 370 375 380
 Pro Arg Pro Ala Ala Gly Pro Pro Arg Ser Pro Ser Gly Pro Arg Gly
 385 390 395 400
 Pro Arg Arg Pro Leu Cys Gly Pro Ser Pro Cys Pro Gly Arg Ala Gly
 405 410 415
 Ser Pro Pro

<210> 199
 <211> 585
 <212> PRT
 <213> HSV2

<400> 199
 Met Asp Pro Tyr Tyr Pro Phe Asp Ala Leu Asp Val Trp Glu His Arg
 5 10 15
 Arg Phe Ile Val Ala Asp Ser Arg Ser Phe Ile Thr Pro Glu Phe Pro
 20 25 30

Arg Asp Phe Trp Met Leu Pro Val Phe Asn Ile Pro Arg Glu Thr Ala
 35 40 45
 Ala Glu Arg Ala Ala Val Leu Gln Ala Gln Arg Thr Ala Ala Ala Ala
 50 55 60
 Ala Leu Glu Asn Ala Ala Leu Gln Ala Ala Glu Leu Pro Val Asp Ile
 65 70 75 80
 Glu Arg Arg Ile Arg Pro Ile Glu Gln Gln Val His His Ile Ala Asp
 85 90 95
 Ala Leu Glu Ala Leu Glu Thr Ala Ala Ala Ala Glu Glu Ala Asp
 100 105 110
 Ala Ala Arg Asp Ala Glu Ala Arg Gly Glu Gly Ala Ala Asp Gly Ala
 115 120 125
 Ala Pro Ser Pro Thr Ala Gly Pro Ala Ala Ala Glu Met Glu Val Gln
 130 135 140
 Ile Val Arg Asn Asp Pro Pro Leu Arg Tyr Asp Thr Asn Leu Pro Val
 145 150 155 160
 Asp Leu Leu His Met Val Tyr Ala Gly Arg Gly Ala Ala Gly Ser Ser
 165 170 175
 Gly Val Val Phe Gly Thr Trp Tyr Arg Thr Ile Gln Glu Arg Thr Ile
 180 185 190
 Ala Asp Phe Pro Leu Thr Thr Arg Ser Ala Asp Phe Arg Asp Gly Arg
 195 200 205
 Met Ser Lys Thr Phe Met Thr Ala Leu Val Leu Ser Leu Gln Ser Cys
 210 215 220
 Gly Arg Leu Tyr Val Gly Gln Arg His Tyr Ser Ala Phe Glu Cys Ala
 225 230 235 240
 Val Leu Cys Leu Tyr Leu Leu Tyr Arg Thr Thr His Glu Ser Ser Pro
 245 250 255
 Asp Arg Asp Arg Ala Pro Val Ala Phe Gly Asp Leu Leu Ala Arg Leu
 260 265 270
 Pro Arg Tyr Leu Ala Arg Leu Ala Ala Val Ile Gly Asp Glu Ser Gly
 275 280 285
 Arg Pro Gln Tyr Arg Tyr Arg Asp Asp Lys Leu Pro Lys Ala Gln Phe
 290 295 300
 Ala Ala Ala Gly Gly Arg Tyr Glu His Gly Ala Leu Ala Thr His Val
 305 310 315 320
 Val Ile Ala Thr Leu Val Arg His Gly Val Leu Pro Ala Ala Pro Gly
 325 330 335
 Asp Val Pro Arg Asp Thr Ser Thr Arg Val Asn Pro Asp Asp Val Ala

Phe Ile Pro Gln Tyr Leu Ser Val Ala
580 585

Pro Met Gly Tyr Val Tyr Gly Arg Ala Cys Pro Ala Glu Gly Leu Glu

20					25					30					
Leu	Leu	Ser	Leu	Leu	Ser	Ala	Arg	Ser	Gly	Asp	Ala	Asp	Val	Ala	Val
		35					40					45			
Ala	Pro	Leu	Ile	Val	Gly	Leu	Thr	Val	Glu	Ser	Gly	Phe	Glu	Ala	Asn
	50					55					60				
Val	Ala	Ala	Val	Val	Gly	Ser	Arg	Thr	Thr	Gly	Leu	Gly	Gly	Thr	Ala
	65					70					75				80
Val	Ser	Leu	Lys	Leu	Met	Pro	Ser	His	Tyr	Ser	Pro	Ser	Val	Tyr	Val
				85					90					95	
Phe	His	Gly	Gly	Arg	His	Leu	Ala	Pro	Ser	Thr	Gln	Ala	Pro	Asn	Leu
			100					105					110		
Thr	Arg	Leu	Cys	Glu	Arg	Ala	Arg	Pro	His	Phe	Gly	Phe	Ala	Asp	Tyr
		115					120					125			
Ala	Pro	Arg	Pro	Cys	Asp	Leu	Lys	His	Glu	Thr	Thr	Gly	Asp	Ala	Leu
						135						140			
Cys	Glu	Arg	Leu	Gly	Leu	Asp	Pro	Asp	Arg	Ala	Leu	Leu	Tyr	Leu	Val
	145					150					155				160
Ile	Thr	Glu	Gly	Phe	Arg	Glu	Ala	Val	Cys	Ile	Ser	Asn	Thr	Phe	Leu
				165					170					175	
His	Leu	Gly	Gly	Met	Asp	Lys	Val	Thr	Ile	Gly	Asp	Ala	Glu	Val	His
			180					185					190		
Arg	Ile	Pro	Val	Tyr	Pro	Leu	Gln	Met	Phe	Met	Pro	Asp	Phe	Ser	Arg
				195			200					205			
Val	Ile	Ala	Asp	Pro	Phe	Asn	Cys	Asn	His	Arg	Ser	Ile	Gly	Glu	Asn
				210			215					220			
Phe	Asn	Tyr	Pro	Leu	Pro	Phe	Phe	Asn	Arg	Pro	Leu	Ala	Arg	Leu	Leu
	225					230					235				240
Phe	Glu	Ala	Val	Val	Gly	Pro	Ala	Ala	Val	Ala	Leu	Arg	Ala	Arg	Asn
				245					250					255	
Val	Asp	Ala	Val	Ala	Arg	Ala	Ala	Ala	His	Leu	Ala	Phe	Asp	Glu	Asn
				260				265					270		
His	Glu	Gly	Ala	Ala	Leu	Pro	Ala	Asp	Ile	Thr	Phe	Thr	Ala	Phe	Glu
				275			280						285		
Ala	Ser	Gln	Gly	Lys	Pro	Gln	Arg	Gly	Ala	Arg	Asp	Ala	Gly	Asn	Lys
				290			295					300			
Gly	Pro	Ala	Gly	Gly	Phe	Glu	Gln	Arg	Leu	Ala	Ser	Val	Met	Ala	Gly
				305			310					315			320
Asp	Ala	Ala	Leu	Ala	Leu	Glu	Ser	Ile	Val	Ser	Met	Ala	Val	Phe	Asp
				325					330					335	

Glu Pro Pro Pro Asp Ile Thr Thr Trp Pro Leu Leu Glu Gly Gln Glu
 340 345 350
 Thr Pro Ala Ala Arg Ala Gly Ala Val Gly Ala Tyr Leu Ala Arg Ala
 355 360 365
 Ala Gly Leu Val Gly Ala Met Val Phe Ser Thr Asn Ser Ala Leu His
 370 375 380
 Leu Thr Glu Val Asp Asp Ala Gly Pro Ala Asp Pro Lys Asp His Ser
 385 390 395 400
 Lys Pro Ser Phe Tyr Arg Phe Phe Leu Val Pro Gly Thr His Val Ala
 405 410 415
 Ala Asn Pro Gln Leu Asp Arg Glu Gly His Val Val Pro Gly Tyr Glu
 420 425 430
 Gly Arg Pro Thr Ala Pro Leu Val Gly Gly Thr Gln Glu Phe Ala Gly
 435 440 445
 Glu His Leu Ala Met Leu Cys Gly Phe Ser Pro Ala Leu Leu Ala Lys
 450 455 460
 Met Leu Phe Tyr Leu Glu Arg Cys Asp Gly Gly Val Ile Val Gly Arg
 465 470 475 480
 Gln Glu Met Asp Val Phe Arg Tyr Val Ala Asp Ser Gly Gln Thr Asp
 485 490 495
 Val Pro Cys Asn Leu Cys Thr Phe Glu Thr Arg His Ala Cys Ala His
 500 505 510
 Thr Thr Leu Met Arg Leu Arg Ala Arg His Pro Lys Phe Ala Ser Ala
 515 520 525
 Ala Arg Gly Ala Ile Gly Val Phe Gly Thr Met Asn Ser Ala Tyr Ser
 530 535 540
 Asp Cys Asp Val Leu Gly Asn Tyr Ala Ala Phe Ser Ala Leu Lys Arg
 545 550 555 560
 Ala Asp Gly Ser Glu Asn Thr Arg Thr Ile Met Gln Glu Thr Tyr Arg
 565 570 575
 Ala Ala Thr Glu Arg Val Met Ala Glu Leu Glu Ala Leu Gln Tyr Val
 580 585 590
 Asp Gln Ala Val Pro Thr Ala Leu Gly Arg Leu Glu Thr Ile Ile Gly
 595 600 605
 Asn Arg Glu Ala Leu His Thr Val Val Asn Asn Ile Lys Gln Leu Val
 610 615 620
 Asp Arg Glu Val Glu Gln Leu Met Arg Asn Leu Ile Glu Gly Arg Asn
 625 630 635 640

Phe Lys Phe Arg Asp Gly Leu Ala Glu Ala Asn His Ala Met Ser Leu
 645 650 655
 Ser Leu Asp Pro Tyr Thr Cys Gly Pro Cys Pro Leu Leu Gln Leu Leu
 660 665 670
 Ala Arg Arg Ser Asn Leu Ala Val Tyr Gln Asp Leu Ala Leu Ser Gln
 675 680 685
 Cys His Gly Val Phe Ala Gly Gln Ser Val Glu Gly Arg Asn Phe Arg
 690 695 700
 Asn Gln Phe Gln Pro Val Leu Arg Arg Arg Val Met Asp Leu Phe Asn
 705 710 715 720
 Asn Gly Phe Leu Ser Ala Lys Thr Leu Thr Val Ala Leu Ser Glu Gly
 725 730 735
 Ala Ala Ile Cys Ala Pro Ser Leu Thr Ala Gly Gln Thr Ala Pro Ala
 740 745 750
 Glu Ser Ser Phe Glu Gly Asp Val Ala Arg Val Thr Leu Gly Phe Pro
 755 760 765
 Lys Glu Leu Arg Val Lys Ser Arg Val Leu Phe Ala Gly Ala Ser Ala
 770 775 780
 Asn Ala Ser Glu Ala Ala Lys Ala Arg Val Ala Ser Leu Gln Ser Ala
 785 790 795 800
 Tyr Gln Lys Pro Asp Lys Arg Val Asp Ile Leu Leu Gly Pro Leu Gly
 805 810 815
 Phe Leu Leu Lys Gln Phe His Ala Val Ile Phe Pro Asn Gly Lys Pro
 820 825 830
 Pro Gly Ser Asn Gln Pro Asn Pro Gln Trp Phe Trp Thr Ala Leu Gln
 835 840 845
 Arg Asn Gln Leu Pro Ala Arg Leu Leu Ser Arg Glu Asp Ile Glu Thr
 850 855 860
 Ile Ala Phe Ile Lys Arg Phe Ser Leu Asp Tyr Gly Ala Ile Asn Phe
 865 870 875 880
 Ile Asn Leu Ala Pro Asn Asn Val Ser Glu Leu Ala Met Tyr Tyr Met
 885 890 895
 Ala Asn Gln Ile Leu Arg Tyr Cys Asp His Ser Thr Tyr Phe Ile Asn
 900 905 910
 Thr Leu Thr Ala Val Ile Ala Gly Ser Arg Arg Pro Pro Ser Val Gln
 915 920 925
 Ala Ala Ala Ala Trp Ala Pro Gln Gly Gly Ala Gly Leu Glu Ala Gly
 930 935 940
 Ala Arg Ala Leu Met Asp Ser Leu Asp Ala His Pro Gly Ala Trp Thr

<400> 201
Met. Ile Pro Ala Ala Leu Pro His Pro Thr Met Lys Arg Gln Gly Asp
 5 10 15

Arg Asp Ile Val Val Thr Gly Val Arg Asn Gln Phe Ala Thr Asp Leu

20					25					30					
Glu	Pro	Gly	Gly	Ser	Val	Ser	Cys	Met	Arg	Ser	Ser	Leu	Ser	Phe	Leu
		35					40					45			
Ser	Leu	Leu	Phe	Asp	Val	Gly	Pro	Arg	Asp	Val	Leu	Ser	Ala	Glu	Ala
	50					55					60				
Ile	Glu	Gly	Cys	Leu	Val	Glu	Gly	Gly	Glu	Trp	Thr	Arg	Ala	Ala	Ala
	65					70					75				80
Gly	Ser	Gly	Pro	Pro	Arg	Met	Cys	Ser	Ile	Ile	Glu	Leu	Pro	Asn	Phe
				85					90					95	
Leu	Glu	Tyr	Pro	Ala	Ala	Arg	Gly	Gly	Leu	Arg	Cys	Val	Phe	Ser	Arg
			100					105					110		
Val	Tyr	Gly	Glu	Val	Gly	Phe	Phe	Gly	Glu	Pro	Thr	Ala	Gly	Leu	Leu
		115						120				125			
Glu	Thr	Gln	Cys	Pro	Ala	His	Thr	Phe	Phe	Ala	Gly	Pro	Trp	Ala	Met
	130					135					140				
Arg	Pro	Leu	Ser	Tyr	Thr	Leu	Leu	Thr	Ile	Gly	Pro	Leu	Gly	Met	Gly
	145					150					155			160	
Leu	Tyr	Arg	Asp	Gly	Asp	Thr	Ala	Tyr	Leu	Phe	Asp	Pro	His	Gly	Leu
				165					170					175	
Pro	Ala	Gly	Thr	Pro	Ala	Phe	Ile	Ala	Lys	Val	Arg	Ala	Gly	Asp	Val
			180					185					190		
Tyr	Pro	Tyr	Leu	Thr	Tyr	Tyr	Ala	His	Asp	Arg	Pro	Lys	Val	Arg	Trp
		195					200					205			
Ala	Gly	Ala	Met	Val	Phe	Phe	Val	Pro	Ser	Gly	Pro	Gly	Ala	Val	Ala
	210					215					220				
Pro	Ala	Asp	Leu	Thr	Ala	Ala	Ala	Leu	His	Leu	Tyr	Gly	Ala	Ser	Glu
	225					230					235			240	
Thr	Tyr	Leu	Gln	Asp	Glu	Pro	Phe	Val	Glu	Arg	Arg	Val	Ala	Ile	Thr
			245						250					255	
His	Pro	Leu	Arg	Gly	Glu	Ile	Gly	Gly	Leu	Gly	Ala	Leu	Phe	Val	Gly
			260					265					270		
Val	Val	Pro	Arg	Gly	Asp	Gly	Glu	Gly	Ser	Gly	Pro	Val	Val	Pro	Ala
		275					280					285			
Leu	Pro	Ala	Pro	Thr	His	Val	Gln	Thr	Pro	Gly	Ala	Asp	Arg	Pro	Pro
	290					295					300				
Glu	Ala	Pro	Arg	Gly	Ala	Ser	Gly	Pro	Pro	Asp	Thr	Pro	Gln	Ala	Gly
	305					310					315			320	
His	Pro	Asn	Arg	Pro	Pro	Asp	Asp	Val	Trp	Ala	Ala	Ala	Leu	Glu	Gly
			325						330					335	

Thr Pro Pro Ala Lys Pro Ser Ala Pro Asp Ala Ala Ala Ser Gly Pro
 340 345 350
 Pro His Ala Ala Pro Pro Pro Gln Thr Pro Ala Gly Asp Ala Ala Glu
 355 360 365
 Glu Ala Glu Asp Leu Arg Val Leu Glu Val Gly Ala Val Pro Val Gly
 370 375 380
 Arg His Arg Ala Arg Tyr Ser Thr Gly Leu Pro Lys Arg Arg Arg Pro
 385 390 395 400
 Thr Trp Thr Pro Pro Ser Ser Val Glu Asp Leu Thr Ser Gly Glu Arg
 405 410 415
 Pro Ala Pro Lys Ala Pro Pro Ala Lys Ala Lys Lys Lys Ser Ala Pro
 420 425 430
 Lys Lys Lys Ala Pro Val Ala Ala Glu Val Pro Ala Ser Ser Pro Thr
 435 440 445
 Pro Ile Ala Ala Thr Val Pro Pro Ala Pro Asp Thr Pro Pro Gln Ser
 450 455 460
 Gly Gln Gly Gly Gly Asp Asp Gly Pro Ala Ser Pro Ser Ser Pro Ser
 465 470 475 480
 Val Leu Glu Thr Leu Gly Ala Arg Arg Pro Pro Glu Pro Pro Gly Ala
 485 490 495
 Asp Leu Ala Gln Leu Phe Glu Val His Pro Asn Val Ala Ala Thr Ala
 500 505 510
 Val Arg Leu Ala Ala Arg Asp Ala Ala Leu Ala Arg Glu Val Ala Ala
 515 520 525
 Cys Ser Gln Leu Thr Ile Asn Ala Leu Arg Ser Pro Tyr Pro Ala His
 530 535 540
 Pro Gly Leu Leu Glu Leu Cys Val Ile Phe Phe Phe Glu Arg Val Leu
 545 550 555 560
 Ala Phe Leu Ile Glu Asn Gly Ala Arg Thr His Thr Gln Ala Gly Val
 565 570 575
 Ala Gly Pro Ala Ala Ala Leu Leu Asp Phe Thr Leu Arg Met Leu Pro
 580 585 590
 Arg Lys Thr Ala Val Gly Asp Phe Leu Ala Ser Thr Arg Met Ser Leu
 595 600 605
 Ala Asp Val Ala Ala His Arg Pro Leu Ile Gln His Val Leu Asp Glu
 610 615 620
 Asn Ser Gln Ile Gly Arg Leu Ala Leu Ala Lys Leu Val Leu Val Ala
 625 630 635 640

Arg Asp Val Ile Arg Glu Thr Asp Ala Phe Tyr Gly Asp Leu Ala Asp
 645 650 655
 Leu Asp Leu Gln Leu Arg Ala Ala Pro Pro Ala Asn Leu Tyr Ala Arg
 660 665 670
 Leu Gly Glu Trp Leu Leu Glu Arg Ser Arg Ala His Pro Asn Thr Leu
 675 680 685
 Phe Ala Pro Ala Thr Pro Thr His Pro Glu Pro Leu Leu His Arg Ile
 690 695 700
 Gln Ala Leu Ala Gln Phe Ala Arg Gly Glu Glu Met Arg Val Glu Ala
 705 710 715 720
 Glu Ala Arg Glu Met Arg Glu Ala Leu Asp Ala Leu Ala Arg Gly Val
 725 730 735
 Asp Ser Val Ser Gln Arg Ala Gly Pro Leu Thr Val Met Pro Val Pro
 740 745 750
 Ala Ala Pro Gly Ala Gly Gly Arg Ala Pro Cys Pro Pro Ala Leu Gly
 755 760 765
 Pro Glu Ala Ile Gln Ala Arg Leu Glu Asp Val Arg Ile Gln Ala Arg
 770 775 780
 Arg Ala Ile Glu Ser Ala Val Lys Glu Tyr Phe His Arg Gly Ala Val
 785 790 795 800
 Tyr Ser Ala Lys Ala Leu Gln Ala Ser Asp Ser His Asp Cys Arg Phe
 805 810 815
 His Val Ala Ser Ala Ala Val Val Pro Met Val Gln Leu Leu Glu Ser
 820 825 830
 Leu Pro Ala Phe Asp Gln His Thr Arg Asp Val Ala Gln Arg Ala Ala
 835 840 845
 Leu Pro Pro Pro Pro Pro Leu Ala Thr Ser Pro Gln Ala Ile Leu Leu
 850 855 860
 Arg Asp Leu Leu Gln Arg Gly Gln Pro Leu Asp Ala Pro Glu Asp Leu
 865 870 875 880
 Ala Ala Trp Leu Ser Val Leu Thr Asp Ala Ala Thr Gln Gly Leu Ile
 885 890 895
 Glu Arg Lys Pro Leu Glu Glu Leu Ala Arg Ser Ile His Gly Ile Asn
 900 905 910
 Asp Gln Gln Ala Arg Arg Ser Ser Gly Leu Ala Glu Leu Gln Arg Phe
 915 920 925
 Asp Ala Leu Asp Ala Ala Leu Ala Gln Gln Leu Asp Ser Asp Ala Ala
 930 935 940
 Phe Val Pro Ala Thr Gly Pro Ala Pro Tyr Val Asp Gly Gly Gly Leu

945	950	955	960
Ser Pro Glu Ala Thr Arg Met Ala Glu Asp Ala Leu Arg Gln Ala Arg	965	970	975
Ala Met Glu Ala Ala Lys Met Thr Ala Glu Leu Ala Pro Glu Ala Arg	980	985	990
Ser Arg Leu Arg Glu Arg Ala His Ala Leu Glu Ala Met Leu Asn Asp	995	1000	1005
Ala Arg Glu Arg Ala Lys Val Ala His Asp Ala Arg Glu Lys Phe Leu	1010	1015	1020
His Lys Leu Gln Gly Val Leu Arg Pro Leu Pro Asp Phe Val Gly Leu	1025	1030	1035
Lys Ala Cys Pro Ala Val Leu Ala Thr Leu Arg Ala Ser Leu Pro Ala	1045	1050	1055
Gly Trp Thr Asp Leu Ala Asp Ala Val Arg Gly Pro Pro Pro Glu Val	1060	1065	1070
Thr Ala Ala Leu Arg Ala Asp Leu Trp Gly Leu Leu Gly Gln Tyr Arg	1075	1080	1085
Glu Ala Leu Glu His Pro Thr Pro Asp Thr Ala Thr Ala Leu Ala Gly	1090	1095	1100
Leu His Pro Ala Phe Val Val Val Leu Lys Thr Leu Phe Ala Asp Ala	1105	1110	1115
Pro Glu Thr Pro Val Leu Val Gln Phe Phe Ser Asp His Ala Pro Thr	1125	1130	1135
Ile Ala Lys Ala Val Ser Asn Ala Ile Asn Ala Gly Ser Ala Ala Val	1140	1145	1150
Ala Thr Ala Ser Pro Ala Ala Thr Val Asp Ala Ala Val Arg Ala His	1155	1160	1165
Gly Ala Leu Ala Asp Ala Val Ser Ala Leu Gly Ala Ala Ala Arg Asp	1170	1175	1180
Pro Ala Ser Pro Leu Ser Phe Leu Ala Val Leu Ala Asp Ser Ala Ala	1185	1190	1195
Gly Tyr Val Lys Ala Thr Arg Leu Ala Leu Glu Ala Arg Gly Ala Ile	1205	1210	1215
Asp Glu Leu Thr Thr Leu Gly Ser Ala Ala Ala Asp Leu Val Val Gln	1220	1225	1230
Ala Arg Arg Ala Cys Ala Gln Pro Glu Gly Asp His Ala Ala Leu Ile	1235	1240	1245
Asp Ala Ala Ala Arg Ala Thr Thr Ala Ala Arg Glu Ser Leu Ala Gly	1250	1255	1260

His Glu Ala Gly Phe Gly Gly Leu Leu His Ala Glu Gly Thr Ala Gly
 1265 1270 1275 1280
 Asp His Ser Pro Ser Gly Arg Ala Leu Gln Glu Leu Gly Lys Val Ile
 1285 1290 1295
 Gly Ala Thr Arg Arg Arg Ala Asp Glu Leu Glu Ala Ala Val Ala Asp
 1300 1305 1310
 Leu Thr Ala Lys Met Ala Ala Gln Arg Ala Arg Gly Ser Ser Glu Arg
 1315 1320 1325
 Trp Ala Ala Gly Val Glu Ala Ala Leu Asp Arg Val Glu Asn Arg Ala
 1330 1335 1340
 Glu Phe Asp Val Val Glu Leu Arg Arg Leu Gln Ala Leu Ala Gly Thr
 1345 1350 1355 1360
 His Gly Tyr Asn Pro Arg Asp Phe Arg Lys Arg Ala Glu Gln Ala Leu
 1365 1370 1375
 Ala Ala Asn Ala Glu Ala Val Thr Leu Ala Leu Asp Thr Ala Phe Ala
 1380 1385 1390
 Phe Asn Pro Tyr Thr Pro Glu Asn Gln Arg His Pro Met Leu Pro Pro
 1395 1400 1405
 Leu Ala Ala Ile His Arg Leu Gly Trp Ser Ala Ala Phe His Ala Ala
 1410 1415 1420
 Ala Glu Thr Tyr Ala Asp Met Phe Arg Val Asp Ala Glu Pro Leu Ala
 1425 1430 1435 1440
 Arg Leu Leu Arg Ile Ala Glu Gly Leu Leu Glu Met Ala Gln Ala Gly
 1445 1450 1455
 Asp Gly Phe Ile Asp Tyr His Glu Ala Val Gly Arg Leu Ala Asp Asp
 1460 1465 1470
 Met Thr Ser Val Pro Gly Leu Arg Arg Tyr Val Pro Phe Phe Gln His
 1475 1480 1485
 Gly Tyr Ala Asp Tyr Val Glu Leu Arg Asp Arg Leu Asp Ala Ile Arg
 1490 1495 1500
 Ala Asp Val His Arg Ala Leu Gly Gly Val Pro Leu Asp Leu Ala Ala
 1505 1510 1515 1520
 Ala Ala Glu Gln Ile Ser Ala Ala Arg Asn Asp Pro Glu Ala Thr Ala
 1525 1530 1535
 Glu Leu Val Arg Thr Gly Val Thr Leu Pro Cys Pro Ser Glu Asp Ala
 1540 1545 1550
 Leu Val Ala Cys Ala Ala Ala Leu Glu Arg Val Asp Gln Ser Pro Val
 1555 1560 1565

Lys Asn Thr Ala Tyr Ala Glu Tyr Val Ala Phe Val Thr Arg Gln Asp
 1570 1575 1580
 Thr Ala Glu Thr Lys Asp Ala Val Val Arg Ala Lys Gln Gln Arg Ala
 1585 1590 1595 1600
 Glu Ala Thr Glu Arg Val Met Ala Gly Leu Arg Glu Ala Leu Ala Ala
 1605 1610 1615
 Arg Glu Arg Arg Ala Gln Ile Glu Ala Glu Gly Leu Ala Asn Leu Lys
 1620 1625 1630
 Thr Met Leu Lys Val Val Ala Val Pro Ala Thr Val Ala Lys Thr Leu
 1635 1640 1645
 Asp Gln Ala Arg Ser Val Ala Glu Ile Ala Asp Gln Val Glu Val Leu
 1650 1655 1660
 Leu Asp Gln Thr Glu Lys Thr Arg Glu Leu Asp Val Pro Ala Val Ile
 1665 1670 1675 1680
 Trp Leu Glu His Ala Gln Arg Thr Phe Glu Thr His Pro Leu Ser Ala
 1685 1690 1695
 Ala Arg Gly Asp Gly Pro Gly Pro Leu Ala Arg His Ala Gly Arg Leu
 1700 1705 1710
 Gly Ala Leu Phe Asp Thr Arg Arg Arg Val Asp Ala Leu Arg Arg Ser
 1715 1720 1725
 Leu Glu Glu Ala Glu Ala Glu Trp Asp Glu Val Trp Gly Arg Phe Gly
 1730 1735 1740
 Arg Val Arg Gly Gly Ala Trp Lys Ser Pro Glu Gly Phe Arg Ala Met
 1745 1750 1755 1760
 His Glu Gln Leu Arg Ala Leu Gln Asp Thr Thr Asn Thr Val Ser Gly
 1765 1770 1775
 Leu Arg Ala Gln Pro Ala Tyr Glu Arg Leu Ser Ala Arg Tyr Gln Gly
 1780 1785 1790
 Val Leu Gly Ala Lys Gly Ala Glu Arg Ala Glu Ala Val Glu Glu Leu
 1795 1800 1805
 Gly Ala Arg Val Thr Lys His Thr Ala Leu Cys Ala Arg Leu Arg Asp
 1810 1815 1820
 Glu Val Val Arg Arg Val Pro Trp Glu Met Asn Phe Asp Ala Leu Gly
 1825 1830 1835 1840
 Gly Leu Leu Ala Glu Phe Asp Ala Ala Ala Ala Asp Leu Ala Pro Trp
 1845 1850 1855
 Ala Val Glu Glu Phe Arg Gly Ala Arg Glu Leu Ile Gln Tyr Arg Met
 1860 1865 1870
 Gly Leu Tyr Ser Ala Tyr Ala Arg Ala Gly Gly Gln Thr Gly Ala Gly

1875	1880	1885
Ala Glu Ser Ala Pro Ala Pro Leu Leu Val Asp Leu Arg Ala Leu Asp 1890 1895 1900		
Ala Arg Ala Arg Ala Ser Ser Ser Pro Glu Gly His Glu Val Asp Pro 1905 1910 1915 1920		
Gln Leu Leu Arg Arg Arg Gly Glu Ala Tyr Leu Arg Ala Gly Gly Asp 1925 1930 1935		
Pro Gly Pro Leu Val Leu Arg Glu Ala Val Ser Ala Leu Asp Leu Pro 1940 1945 1950		
Phe Ala Thr Ser Phe Leu Ala Pro Asp Gly Thr Pro Leu Gln Tyr Ala 1955 1960 1965		
Leu Cys Phe Pro Ala Val Thr Asp Lys Leu Gly Ala Leu Leu Met Arg 1970 1975 1980		
Pro Glu Ala Ala Cys Val Arg Pro Pro Leu Pro Thr Asp Val Leu Glu 1985 1990 1995 2000		
Ser Ala Pro Thr Val Thr Ala Met Tyr Val Leu Thr Val Val Asn Arg 2005 2010 2015		
Leu Gln Leu Ala Leu Ser Asp Ala Gln Ala Ala Asn Phe Gln Leu Phe 2020 2025 2030		
Gly Arg Phe Val Arg His Arg Gln Ala Thr Trp Gly Ala Ser Met Asp 2035 2040 2045		
Ala Ala Ala Glu Leu Tyr Val Ala Leu Val Ala Thr Thr Leu Thr Arg 2050 2055 2060		
Glu Phe Gly Cys Arg Trp Ala Gln Leu Gly Trp Ala Ser Gly Ala Ala 2065 2070 2075 2080		
Ala Pro Arg Pro Pro Pro Gly Pro Arg Gly Ser Gln Arg His Cys Val 2085 2090 2095		
Ala Phe Asn Glu Asn Asp Val Leu Val Ala Leu Val Ala Gly Val Pro 2100 2105 2110		
Glu His Ile Tyr Asn Phe Trp Arg Leu Asp Leu Val Arg Gln His Glu 2115 2120 2125		
Tyr Met His Leu Thr Leu Glu Arg Ala Phe Glu Asp Ala Ala Glu Ser 2130 2135 2140		
Met Leu Phe Val Gln Arg Leu Thr Pro His Pro Asp Ala Arg Ile Arg 2145 2150 2155 2160		
Val Leu Pro Thr Phe Leu Asp Gly Gly Pro Pro Thr Arg Gly Leu Leu 2165 2170 2175		
Phe Gly Thr Arg Leu Ala Asp Trp Arg Arg Gly Lys Leu Ser Glu Thr 2180 2185 2190		

Asp Pro Leu Ala Pro Trp Arg Ser Ala Leu Glu Leu Gly Thr Gln Arg
 2195 2200 2205
 Arg Asp Val Pro Ala Leu Gly Lys Leu Ser Pro Ala Gln Ala Leu Ala
 2210 2215 2220
 Ala Val Ser Val Leu Gly Arg Met Cys Leu Pro Ser Ala Ala Leu Ala
 2225 2230 2235 2240
 Ala Leu Trp Thr Cys Met Phe Pro Asp Asp Tyr Thr Glu Tyr Asp Ser
 2245 2250 2255
 Phe Asp Ala Leu Leu Ala Ala Arg Leu Glu Ser Gly Gln Thr Leu Gly
 2260 2265 2270
 Pro Ala Gly Gly Arg Glu Ala Ser Leu Pro Glu Ala Pro His Ala Leu
 2275 2280 2285
 Tyr Arg Pro Thr Gly Gln His Val Ala Val Leu Ala Ala Ala Thr His
 2290 2295 2300
 Arg Thr Pro Ala Ala Arg Val Thr Ala Met Asp Leu Val Leu Ala Ala
 2305 2310 2315 2320
 Val Leu Leu Gly Ala Pro Val Val Val Ala Leu Arg Asn Thr Thr Ala
 2325 2330 2335
 Phe Ser Arg Glu Ser Glu Leu Glu Leu Cys Leu Thr Leu Phe Asp Ser
 2340 2345 2350
 Arg Pro Gly Gly Pro Asp Ala Ala Leu Arg Asp Val Val Ser Ser Asp
 2355 2360 2365
 Ile Glu Thr Trp Ala Val Gly Leu Leu His Thr Asp Leu Asn Pro Ile
 2370 2375 2380
 Glu Asn Ala Cys Leu Ala Ala Gln Leu Pro Arg Leu Ser Ala Leu Ile
 2385 2390 2395 2400
 Ala Glu Arg Pro Leu Ala Asp Gly Pro Pro Cys Leu Val Leu Val Asp
 2405 2410 2415
 Ile Ser Met Thr Pro Val Ala Val Leu Trp Glu Ala Pro Glu Pro Pro
 2420 2425 2430
 Gly Pro Pro Asp Val Arg Phe Val Gly Ser Glu Ala Thr Glu Glu Leu
 2435 2440 2445
 Pro Phe Val Ala Thr Ala Gly Asp Val Leu Ala Ala Ser Ala Ala Asp
 2450 2455 2460
 Ala Asp Pro Phe Phe Ala Arg Ala Ile Leu Gly Arg Pro Phe Asp Ala
 2465 2470 2475 2480
 Ser Leu Leu Thr Gly Glu Leu Phe Pro Gly His Pro Val Tyr Gln Arg
 2485 2490 2495

Pro Leu Ala Asp Glu Ala Gly Pro Ser Ala Pro Thr Ala Ala Arg Asp
 2500 2505 2510
 Pro Arg Asp Leu Ala Gly Gly Asp Gly Gly Ser Gly Pro Glu Asp Pro
 2515 2520 2525
 Ala Ala Pro Pro Ala Arg Gln Ala Asp Pro Gly Val Leu Ala Pro Thr
 2530 2535 2540
 Leu Leu Thr Asp Ala Thr Thr Gly Glu Pro Val Pro Pro Arg Met Trp
 2545 2550 2555 2560
 Ala Trp Ile His Gly Leu Glu Glu Leu Ala Ser Asp Asp Ala Gly Gly
 2565 2570 2575
 Pro Thr Pro Asn Pro Ala Pro Ala Leu Leu Pro Pro Pro Ala Thr Asp
 2580 2585 2590
 Gln Ser Val Pro Thr Ser Gln Tyr Ala Pro Arg Pro Ile Gly Pro Ala
 2595 2600 2605
 Ala Thr Ala Arg Glu Thr Arg Pro Ser Val Pro Pro Gln Gln Asn Thr
 2610 2615 2620
 Gly Arg Val Pro Val Ala Pro Arg Asp Asp Pro Arg Pro Ser Pro Pro
 2625 2630 2635 2640
 Thr Pro Ser Pro Pro Ala Asp Ala Ala Leu Pro Pro Pro Ala Phe Ser
 2645 2650 2655
 Gly Ser Ala Ala Ala Phe Ser Ala Ala Val Pro Arg Val Arg Arg Ser
 2660 2665 2670
 Arg Arg Thr Arg Ala Lys Ser Arg Ala Pro Arg Ala Ser Ala Pro Pro
 2675 2680 2685
 Glu Gly Trp Arg Pro Pro Ala Leu Pro Ala Pro Val Ala Pro Val Ala
 2690 2695 2700
 Ala Ser Ala Arg Pro Pro Asp Gln Pro Pro Thr Pro Glu Ser Ala Pro
 2705 2710 2715 2720
 Pro Ala Trp Val Ser Ala Leu Pro Leu Pro Pro Gly Pro Ala Ser Ala
 2725 2730 2735
 Arg Gly Ala Phe Pro Ala Pro Thr Leu Ala Pro Ile Pro Pro Pro Pro
 2740 2745 2750
 Ala Glu Gly Ala Val Val Pro Gly Gly Asp Arg Arg Arg Gly Arg Arg
 2755 2760 2765
 Gln Thr Thr Ala Gly Pro Ser Pro Thr Pro Pro Arg Gly Pro Ala Ala
 2770 2775 2780
 Gly Pro Pro Arg Arg Leu Thr Arg Pro Ala Val Ala Ser Leu Ser Ala
 2785 2790 2795 2800
 Ser Leu Asn Ser Leu Pro Ser Pro Arg Asp Pro Ala Asp His Ala Ala

2805	2810	2815
Ala Val Ser Ala Ala Ala Ala Val Pro Pro Ser Pro Gly Leu Ala		
2820	2825	2830
Pro Pro Thr Ser Ala Val Gln Thr Ser Pro Pro Pro Leu Ala Pro Gly		
2835	2840	2845
Pro Val Ala Pro Ser Glu Pro Leu Cys Gly Trp Val Val Pro Gly Gly		
2850	2855	2860
Pro Val Ala Arg Arg Pro Pro Pro Gln Ser Pro Ala Thr Lys Pro Ala		
2865	2870	2880
Ala Arg Thr Arg Ile Arg Ala Arg Ser Val Pro Gln Pro Pro Leu Pro		
2885	2890	2895
Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln		
2900	2905	2910
Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro		
2915	2920	2925
Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro		
2930	2935	2940
Leu Pro Pro Val Thr Arg Thr Leu Thr Pro Gln Ser Arg Asp Ser Val		
2945	2950	2955
Pro Thr Pro Glu Ser Pro Thr His Thr Asn Thr His Leu Pro Val Ser		
2965	2970	2975
Ala Val Thr Ser Trp Ala Ser Ser Leu Ala Leu His Val Asp Ser Ala		
2980	2985	2990
Pro Pro Pro Ala Ser Leu Leu Gln Thr Leu His Ile Ser Ser Asp Asp		
2995	3000	3005
Glu His Ser Asp Ala Asp Ser Leu Arg Phe Ser Asp Ser Asp Asp Thr		
3010	3015	3020
Glu Ala Leu Asp Pro Leu Pro Pro Glu Pro His Leu Pro Pro Ala Asp		
3025	3030	3035
Glu Pro Pro Gly Pro Leu Ala Ala Asp His Leu Gln Ser Pro His Ser		
3045	3050	3055
Gln Phe Gly Pro Leu Pro Val Gln Ala Asn Ala Val Leu Ser Arg Arg		
3060	3065	3070
Tyr Val Arg Ser Thr Gly Arg Ser Ala Leu Ala Val Leu Ile Arg Ala		
3075	3080	3085
Cys Arg Arg Ile Gln Gln Gln Leu Gln Arg Thr Arg Arg Ala Leu Phe		
3090	3095	3100
Gln Arg Ser Asn Ala Val Leu Thr Ser Leu His His Val Arg Met Leu		
3105	3110	3115
		3120

```
<210> 202
<211> 1114
<212> PRT
<213> HSV2
```

<400> 202																	
Met	Ser	Asp	Ser	Ala	Leu	Gln	Val	Pro	Ala	Pro	Ala	Gly	Met	Thr	Pro		
				5					10					15			
Pro	Ser	Ala	Pro	Pro	Pro	Asn	Gly	Pro	Leu	Gln	Val	Leu	Leu	Gly	Ser		
			20					25					30				
Leu	Thr	Asn	Leu	Arg	Arg	Pro	Pro	Ser	Pro	Ser	Ser	Glu	Pro	Ala	Gly		
		35					40					45					
Ser	Ala	Asp	Glu	Pro	Ala	Phe	Leu	Ser	Ala	Ala	Lys	Leu	His	Ala	Ala		
	50					55					60						
Thr	Ala	Ala	Phe	Leu	Leu	Ser	Gly	Ala	Ala	Val	Gly	Pro	Ala	Glu	Ala		
	65				70					75					80		
Arg	Ala	Cys	Trp	His	Pro	Leu	Leu	Glu	Gln	Leu	Cys	Ala	Leu	His	Arg		
				85					90					95			
Ala	His	Gly	Leu	Pro	Glu	Thr	Ala	Leu	Leu	Ala	Glu	Asn	Leu	Pro	Gly		
			100					105					110				
Leu	Leu	Val	His	Arg	Met	Ala	Val	Ala	Leu	Pro	Glu	Thr	Pro	Glu	Ala		
		115					120					125					
Ala	Phe	Arg	Glu	Met	Asp	Val	Ile	Lys	Asp	Thr	Val	Leu	Ala	Ile	Thr		
	130					135					140						
Gly	Ser	Asp	Thr	Thr	His	Ala	Leu	Glu	Ala	Ala	Gly	Leu	Arg	Thr	Thr		
145					150					155					160		
Ala	Ala	Leu	Gly	Pro	Val	Arg	Val	Arg	Gln	Cys	Ala	Val	Glu	Trp	Ile		
				165					170					175			
Asp	Arg	Trp	Arg	Thr	Val	Thr	Gln	Ser	Cys	Leu	Ala	Met	Asn	Pro	Arg		
			180					185					190				
Thr	Ser	Leu	Glu	Ala	Leu	Gly	Glu	Met	Ser	Leu	Lys	Met	Ser	Pro	Val		
		195					200					205					
Pro	Leu	Gly	Gln	Pro	Gly	Ala	Asn	Leu	Thr	Thr	Pro	Ala	Tyr	Ser	Leu		
	210					215					220						
Leu	Phe	Pro	Ser	Pro	Ile	Val	Gln	Glu	Gly	Leu	Arg	Phe	Leu	Ala	Leu		
225					230					235					240		
Val	Ser	Asn	Trp	Val	Thr	Leu	Phe	Ser	Ala	His	Leu	Gln	Arg	Ile	Asp		
				245					250					255			

Asp Ala Ala Leu Thr Pro Leu Thr Arg Ala Leu Phe Thr Leu Ala Leu
 260 265 270
 Val Asp Asp Tyr Leu Thr Thr Pro Asp Arg Gly Ala Val Val Pro Pro
 275 280 285
 Pro Leu Leu Ala Gln Phe Gln His Thr Val Arg Glu Ile Asp Pro Ala
 290 295 300
 Ile Met Ile Pro Pro Leu Glu Ala Thr Lys Met Val Arg Ser Arg Glu
 305 310 315 320
 Glu Val Arg Val Ser Thr Ala Leu Ser Arg Val Ser Pro Arg Ser Ala
 325 330 335
 Cys Ala Pro Pro Gly Thr Leu Met Ala Arg Val Arg Thr Asp Ala Ala
 340 345 350
 Val Phe Asp Pro Asp Val Pro Phe Leu Ser Ala Ser Ala Leu Ala Ile
 355 360 365
 Phe Arg Pro Ala Val Thr Gly Leu Leu Gln Leu Gly Glu Pro Pro Ser
 370 375 380
 Ala Gly Ala Gln Gln Arg Leu Leu Ala Leu Leu Gln Gln Thr Trp Ala
 385 390 395 400
 Leu Val Gln Asn Ser Asn Ser Pro Ser Val Val Ile Asn Thr Leu Thr
 405 410 415
 Asp Ala Gly Phe Thr Pro Ala His Cys Thr Gln Tyr Ile Ser Ala Leu
 420 425 430
 Glu Gly Phe Leu Val Ala Gly Val Pro Ala Arg Thr Pro Pro Gly His
 435 440 445
 Gly Leu Ser Glu Ile Gln Gln Leu Phe Gly Cys Ile Ala Leu Ala Gly
 450 455 460
 Ala Asn Val Phe Gly Leu Ala Arg Glu Tyr Gly His Tyr Ala Gly Tyr
 465 470 475 480
 Val Lys Thr Phe Arg Arg Ile Gln Gly Ala Ser Glu His Thr His Gly
 485 490 495
 Arg Leu Cys Glu Ala Val Gly Leu Ser Gly Gly Val Leu Ser Gln Thr
 500 505 510
 Leu Ala Arg Ile Met Gly Pro Ala Val Pro Thr Glu His Leu Ala Ser
 515 520 525
 Leu Arg Arg Thr Leu Val Gly Glu Phe Glu Thr Ala Glu Arg Arg Phe
 530 535 540
 Ser Ala Gly Gln Pro Ser Leu Leu Arg Glu Thr Ala Leu Ile Trp Leu
 545 550 555 560

Asp Val Tyr Gly Gln Thr His Trp Asp Leu Thr Pro Thr Thr Pro Ala
 565 570 575
 Thr Pro Leu Ser Ala Leu Leu Pro Val Gly Pro Pro Ser His Ala Pro
 580 585 590
 Ser Val His Leu Ala Ala Ala Thr Lys Ile Arg Phe Pro Ala Leu Glu
 595 600 605
 Gly Ile His Pro Asn Val Leu Ala Asp Pro Gly Phe Val Pro Tyr Val
 610 615 620
 Leu Ala Leu Val Val Gly Asp Ala Leu Arg Ala Thr Cys Asn Ala Ala
 625 630 635 640
 Tyr Leu Pro Arg Pro Ile Glu Phe Ala Leu Arg Val Leu Ala Trp Ala
 645 650 655
 Arg Asp Phe Gly Leu Gly Tyr Leu Pro Thr Val Glu Gly His Arg Thr
 660 665 670
 Lys Leu Gly Ala Leu Ile Thr Leu Leu Glu Pro Ala Thr Arg Ala Gly
 675 680 685
 Val Gly Pro Thr Met Gln Met Ala Asp Asn Ile Glu Gln Leu Leu Arg
 690 695 700
 Glu Leu Tyr Val Ile Ala Arg Gly Ala Val Glu Gln Leu Arg Pro Ala
 705 710 715 720
 Val Gln Leu Pro Pro Pro Gln Pro Pro Glu Val Gly Ser Ser Leu Leu
 725 730 735
 Leu Ile Ser Met Tyr Ala Leu Ala Ala Arg Gly Val Leu Gln Glu Leu
 740 745 750
 Ala Glu Arg Ala Asp Pro Leu Val Arg Gln Leu Glu Asp Ala Ile Val
 755 760 765
 Leu Leu Arg Leu His Met Arg Thr Leu Ala Ala Phe Phe Glu Cys Arg
 770 775 780
 Phe Glu Ser Asp Gly His Arg Leu Tyr Ala Val Val Ala Asp Ala His
 785 790 795 800
 Glu Arg Leu Gly Pro Trp Arg Pro Glu Ala Met Gly Asp Ala Val Ser
 805 810 815
 Gln Tyr Cys Gly Met Tyr His Asp Ala Lys Arg Ala Leu Val Ala Ser
 820 825 830
 Leu Ala Gly Leu Arg Ser Val Val Thr Glu Thr Thr Ala His Leu Gly
 835 840 845
 Val Cys Asp Glu Leu Ala Ala Gln Val Ser His Glu Gly Asn Val Leu
 850 855 860
 Ala Val Val Arg Arg Glu Ile His Gly Phe Leu Ala Ile Val Ser Gly

865 870 875 880
 Ile His Ala Arg Ala Ser Lys Leu Met Ser Gly Asp Gln Val Pro Gly
 885 890 895
 Phe Cys Tyr Met Ser Gln Phe Leu Ala Arg Trp Arg Arg Leu Ser Ala
 900 905 910
 Gly Tyr Gln Ala Ala Arg Ala Ala Thr Gly Pro Glu Arg Val Ala Glu
 915 920 925
 Phe Val Gln Glu Leu His Asp Thr Trp Lys Gly Leu Gln Thr Glu Arg
 930 935 940
 Ala Leu Val Val Ala Arg Phe Ala Ser Ser Ala Asp Gln Arg Thr Ala
 945 950 955 960
 Ala Ile Gln Glu Val Met Ala His Ala Thr Glu Asp Ala Pro Pro Ser
 965 970 975
 Pro Ala Ala Asp Leu Val Val Leu Thr Asn Arg His Asp Leu Gly Ala
 980 985 990
 Trp Gly Asp Tyr Ser Leu Gly Pro Leu Gly Gln Pro Thr Val Val Pro
 995 1000 1005
 Asp Ser Val Asp Leu Ser Pro Gln Gly Leu Ala Ala Thr Leu Ser Met
 1010 1015 1020
 Asp Trp Leu Leu Ile Asn Glu Leu Leu Gln Val Thr Asp Gly Val Phe
 1025 1030 1035 1040
 Arg Ala Ser Ala Phe Arg Pro Ser Ala Gly Pro Gly Ala Pro Gly Asp
 1045 1050 1055
 Leu Glu Ala Gln Asp Ala Gly Gly Ser Thr Pro Glu Pro Thr Thr Pro
 1060 1065 1070
 Gly Pro Gln Asp Thr Gln Ala Arg Ala Pro Ser Thr Arg Pro Ala Gly
 1075 1080 1085
 Arg Glu Thr Val Pro Trp Pro Asn Thr Pro Val Glu Asp Asp Glu Met
 1090 1095 1100
 Thr Pro Gln Glu Thr Pro Pro Val His Pro
 1105 1110

<210> 203
 <211> 1142
 <212> PRT
 <213> HSV2

<400> 203
 Met Ala Asn Arg Pro Ala Ala Ser Ala Leu Ala Gly Ala Arg Ser Pro
 5 10 15

Ser Glu Arg Gln Glu Pro Arg Glu Pro Glu Val Ala Pro Pro Gly Gly

20	25	30
Asp His Val Phe Cys Arg Lys Val Ser Gly Val Met Val Leu Ser Ser		
35	40	45
Asp Pro Pro Gly Pro Ala Ala Tyr Arg Ile Ser Asp Ser Ser Phe Val		
50	55	60
Gln Cys Gly Ser Asn Cys Ser Met Ile Ile Asp Gly Asp Val Ala Arg		
65	70	75
Gly His Leu Arg Asp Leu Glu Gly Ala Thr Ser Thr Gly Ala Phe Val		
85	90	95
Ala Ile Ser Asn Val Ala Ala Gly Gly Asp Gly Arg Thr Ala Val Val		
100	105	110
Ala Leu Gly Gly Thr Ser Gly Pro Ser Ala Thr Thr Ser Val Gly Thr		
115	120	125
Gln Thr Ser Gly Glu Phe Leu His Gly Asn Pro Arg Thr Pro Glu Pro		
130	135	140
Gln Gly Pro Gln Ala Val Pro Pro Pro Pro Pro Pro Phe Pro Trp		
145	150	155
Gly His Glu Cys Cys Ala Arg Arg Asp Ala Arg Gly Gly Ala Glu Lys		
165	170	175
Asp Val Gly Ala Ala Glu Ser Trp Ser Asp Gly Pro Ser Ser Asp Ser		
180	185	190
Glu Thr Glu Asp Ser Asp Ser Ser Asp Glu Asp Thr Gly Ser Glu Thr		
195	200	205
Leu Ser Arg Ser Ser Ser Ile Trp Ala Ala Gly Ala Thr Asp Asp Asp		
210	215	220
Asp Ser Asp Ser Asp Ser Arg Ser Asp Asp Ser Val Gln Pro Asp Val		
225	230	235
Val Val Arg Arg Arg Trp Ser Asp Gly Pro Ala Pro Val Ala Phe Pro		
245	250	255
Lys Pro Arg Arg Pro Gly Asp Ser Pro Gly Asn Pro Gly Leu Gly Ala		
260	265	270
Gly Thr Gly Pro Gly Ser Ala Thr Asp Pro Arg Ala Ser Ala Asp Ser		
275	280	285
Asp Ser Ala Ala His Ala Ala Ala Pro Gln Ala Asp Val Ala Pro Val		
290	295	300
Leu Asp Ser Gln Pro Thr Val Gly Thr Asp Pro Gly Tyr Pro Val Pro		
305	310	315
Leu Glu Leu Thr Pro Glu Asn Ala Glu Ala Val Ala Arg Phe Leu Gly		
325	330	335

Asp Ala Val Asp Arg Glu Pro Ala Leu Met Leu Glu Tyr Phe Cys Arg
 340 345 350
 Cys Ala Arg Glu Glu Ser Lys Arg Val Pro Pro Arg Thr Phe Gly Ser
 355 360
 Ala Pro Arg Leu Thr Glu Asp Asp Phe Gly Leu Leu Asn Tyr Ala Leu
 370 375 380
 Ala Glu Met Arg Arg Leu Cys Leu Asp Leu Pro Pro Val Pro Pro Asn
 385 390 395 400
 Ala Tyr Thr Pro Tyr His Leu Arg Glu Tyr Ala Thr Arg Leu Val Asn
 405 410 415
 Gly Phe Lys Pro Leu Val Arg Arg Ser Ala Arg Leu Tyr Arg Ile Leu
 420 425 430
 Gly Val Leu Val His Leu Arg Ile Arg Thr Arg Glu Ala Ser Phe Glu
 435 440 445
 Glu Trp Met Arg Ser Lys Glu Val Asp Leu Asp Phe Gly Leu Thr Glu
 450 455 460
 Arg Leu Arg Glu His Glu Ala Gln Leu Met Ile Leu Ala Gln Ala Leu
 465 470 475 480
 Asn Pro Tyr Asp Cys Leu Ile His Ser Thr Pro Asn Thr Leu Val Glu
 485 490 495
 Arg Gly Leu Gln Ser Ala Leu Lys Tyr Glu Glu Phe Tyr Leu Lys Arg
 500 505 510
 Phe Gly Gly His Tyr Met Glu Ser Val Phe Gln Met Tyr Thr Arg Ile
 515 520 525
 Ala Gly Phe Leu Ala Cys Arg Ala Thr Arg Gly Met Arg His Ile Ala
 530 535 540
 Leu Gly Arg Gln Gly Ser Trp Trp Glu Met Phe Lys Phe Phe Phe His
 545 550 555 560
 Arg Leu Tyr Asp His Gln Ile Val Pro Ser Thr Pro Ala Met Leu Asn
 565 570 575
 Leu Gly Thr Arg Asn Tyr Tyr Thr Ser Ser Cys Tyr Leu Val Asn Pro
 580 585 590
 Gln Ala Thr Thr Asn Gln Ala Thr Leu Arg Ala Ile Thr Gly Asn Val
 595 600 605
 Ser Ala Ile Leu Ala Arg Asn Gly Gly Ile Gly Leu Cys Met Gln Ala
 610 615 620
 Phe Asn Asp Ala Ser Pro Gly Thr Ala Ser Ile Met Pro Ala Leu Lys
 625 630 635 640

Val Leu Asp Ser Leu Val Ala Ala His Asn Lys Gln Ser Thr Arg Pro
 645 650 655
 Thr Gly Ala Cys Val Tyr Leu Glu Pro Trp His Ser Asp Val Arg Ala
 660 665 670
 Val Leu Arg Met Lys Gly Val Leu Ala Gly Glu Glu Ala Gln Arg Cys
 675 680 685
 Asp Asn Ile Phe Ser Ala Leu Trp Met Pro Asp Leu Phe Phe Lys Arg
 690 695 700
 Leu Ile Arg His Leu Asp Gly Glu Lys Asn Val Thr Trp Ser Leu Phe
 705 710 715 720
 Asp Arg Asp Thr Ser Met Ser Leu Ala Asp Phe His Gly Glu Glu Phe
 725 730 735
 Glu Lys Leu Tyr Glu His Leu Glu Ala Met Gly Phe Gly Glu Thr Ile
 740 745 750
 Pro Ile Gln Asp Leu Ala Tyr Ala Ile Val Arg Ser Ala Ala Thr Thr
 755 760 765
 Gly Ser Pro Phe Ile Met Phe Lys Asp Ala Val Asn Arg His Tyr Ile
 770 775 780
 Tyr Asp Thr Gln Gly Ala Ala Ile Ala Gly Ser Asn Leu Cys Thr Glu
 785 790 795 800
 Ile Val His Pro Ala Ser Lys Arg Ser Ser Gly Val Cys Asn Leu Gly
 805 810 815
 Ser Val Asn Leu Ala Arg Cys Val Ser Arg Gln Thr Phe Asp Phe Gly
 820 825 830
 Arg Leu Arg Asp Ala Val Gln Ala Cys Val Leu Met Val Asn Ile Met
 835 840 845
 Ile Asp Ser Thr Leu Gln Pro Thr Pro Gln Cys Thr Arg Gly Asn Asp
 850 855 860
 Asn Leu Arg Ser Met Gly Ile Gly Met Gln Gly Leu His Thr Ala Cys
 865 870 875 880
 Leu Lys Met Gly Leu Asp Leu Glu Ser Ala Glu Phe Arg Asp Leu Asn
 885 890 895
 Thr His Ile Ala Glu Val Met Leu Leu Ala Ala Met Lys Thr Ser Asn
 900 905 910
 Ala Leu Cys Val Arg Gly Ala Arg Pro Phe Ser His Phe Lys Arg Ser
 915 920 925
 Met Tyr Arg Ala Gly Arg Phe His Trp Glu Arg Phe Ser Asn Ala Ser
 930 935 940
 Pro Arg Tyr Glu Gly Glu Trp Glu Met Leu Arg Gln Ser Met Met Lys

```
<210> 204
<211> 492
<212> PRT
<213> HSV2
```

<400> 204

Met Gly Leu Phe Gly Met Met Lys Phe Ala Gln Thr His His Leu Val
5 10 15

Lys Arg Arg Gly Leu Arg Ala Pro Glu Gly Tyr Phe Thr Pro Ile Ala
20 25 30

Val Asp Leu Trp Asn Val Met Tyr Thr Leu Val Val Lys Tyr Gln Arg
35 40 45

Arg Tyr Pro Ser Tyr Asp Arg Glu Ala Ile Thr Leu His Cys Leu Cys
50 55 60

Ser Met Leu Arg Val Phe Thr Gln Lys Ser Leu Phe Pro Ile Phe Val

65		70		75		80
Thr Asp Arg Gly Val Glu Cys Thr Glu Pro Val Val Phe Gly Ala Lys	85		90		95	
Ala Ile Leu Ala Arg Thr Thr Ala Gln Cys Arg Thr Asp Glu Glu Ala	100		105		110	
Ser Asp Val Asp Ala Ser Pro Pro Pro Ser Pro Ile Thr Asp Ser Arg	115		120		125	
Pro Ser Phe Ala Phe Ser Asn Met Arg Arg Arg Gly His Ala Phe Ala	130		135		140	
Pro Gly Asp Arg Gly Thr Arg Ala Ala Gly Pro Gly Pro Ala Ala Pro	145		150		155	160
Ser Gly Ala Pro Ser Lys Pro Ala Leu Arg Leu Ala His Leu Phe Cys	165		170		175	
Ile Arg Val Leu Arg Ala Leu Gly Tyr Ala Tyr Ile Asn Ser Gly Gln	180		185		190	
Leu Glu Ala Asp Asp Ala Cys Ala Asn Leu Tyr His Thr Asn Thr Val	195		200		205	
Ala Tyr Val His Thr Thr Asp Thr Asp Leu Leu Leu Met Gly Cys Asp	210		215		220	
Ile Val Leu Asp Ile Ser Thr Gly Tyr Ile Pro Thr Ile His Cys Arg	225		230		235	240
Asp Leu Leu Gln Tyr Phe Lys Met Ser Tyr Pro Gln Phe Leu Ala Leu	245		250		255	
Phe Val Arg Cys His Thr Asp Leu His Pro Asn Asn Thr Tyr Ala Ser	260		265		270	
Val Glu Asp Val Leu Arg Glu Cys His Trp Thr Ala Pro Ser Arg Ser	275		280		285	
Gln Ala Arg Arg Ala Ala Arg Arg Glu Arg Ala Asn Ser Arg Ser Leu	290		295		300	
Glu Ser Met Pro Thr Leu Thr Ala Ala Pro Val Gly Leu Glu Thr Arg	305		310		315	320
Ile Ser Trp Thr Glu Ile Leu Ala Gln Gln Ile Ala Gly Glu Asp Asp	325		330		335	
Tyr Glu Glu Asp Pro Pro Leu Gln Pro Pro Asp Val Ala Gly Gly Pro	340		345		350	
Arg Asp Gly Ala Arg Ser Ser Ser Ser Glu Ile Leu Thr Pro Pro Glu	355		360		365	
Leu Val Gln Val Pro Asn Ala Gln Arg Val Ala Glu His Arg Gly Tyr	370		375		380	

Val Ala Gly Arg Arg Arg His Val Ile His Asp Ala Pro Glu Ala Leu
 385 390 395 400
 Asp Trp Leu Pro Asp Pro Met Thr Ile Ala Glu Leu Val Glu His Arg
 405 410 415
 Tyr Val Lys Tyr Val Ile Ser Leu Ile Ser Pro Lys Glu Arg Gly Pro
 420 425 430
 Trp Thr Leu Leu Lys Arg Leu Pro Ile Tyr Gln Asp Leu Arg Asp Glu
 435 440 445
 Asp Leu Ala Arg Ser Ile Val Thr Arg His Ile Thr Ala Pro Asp Ile
 450 455 460
 Ala Asp Arg Phe Leu Ala Gln Leu Trp Ala His Ala Pro Pro Pro Ala
 465 470 475 480
 Phe Tyr Lys Asp Val Leu Ala Lys Phe Trp Asp Glu
 485 490

<210> 205
 <211> 490
 <212> PRT
 <213> HSV2

<400> 205
 Met Asp Leu Leu Val Asp Asp Leu Phe Ala Asp Ala Asp Gly Val Ser
 5 10 15
 Pro Pro Pro Pro Arg Pro Ala Gly Gly Pro Lys Asn Thr Pro Ala Ala
 20 25 30
 Pro Pro Leu Tyr Ala Thr Gly Arg Leu Ser Gln Ala Gln Leu Met Pro
 35 40 45
 Ser Pro Pro Met Pro Val Pro Pro Ala Ala Leu Phe Asn Arg Leu Leu
 50 55 60
 Asp Asp Leu Gly Phe Ser Ala Gly Pro Ala Leu Cys Thr Met Leu Asp
 65 70 75 80
 Thr Trp Asn Glu Asp Leu Phe Ser Gly Phe Pro Thr Asn Ala Asp Met
 85 90 95
 Tyr Arg Glu Cys Lys Phe Leu Ser Thr Leu Pro Ser Asp Val Ile Asp
 100 105 110
 Trp Gly Asp Ala His Val Pro Glu Arg Ser Pro Ile Asp Ile Arg Ala
 115 120 125
 His Gly Asp Val Ala Phe Pro Thr Leu Pro Ala Thr Arg Asp Glu Leu
 130 135 140
 Pro Ser Tyr Tyr Glu Ala Met Ala Gln Phe Phe Arg Gly Glu Leu Arg
 145 150 155 160

Ala Arg Glu Glu Ser Tyr Arg Thr Val Leu Ala Asn Phe Cys Ser Ala
 165 170 175
 Leu Tyr Arg Tyr Leu Arg Ala Ser Val Arg Gln Leu His Arg Gln Ala
 180 185 190
 His Met Arg Gly Arg Asn Arg Asp Leu Arg Glu Met Leu Arg Thr Thr
 195 200 205
 Ile Ala Asp Arg Tyr Tyr Arg Glu Thr Ala Arg Leu Ala Arg Val Leu
 210 215 220
 Phe Leu His Leu Tyr Leu Phe Leu Ser Arg Glu Ile Leu Trp Ala Ala
 225 230 235 240
 Tyr Ala Glu Gln Met Met Arg Pro Asp Leu Phe Asp Gly Leu Cys Cys
 245 250 255
 Asp Leu Glu Ser Trp Arg Gln Leu Ala Cys Leu Phe Gln Pro Leu Met
 260 265 270
 Phe Ile Asn Gly Ser Leu Thr Val Arg Gly Val Pro Val Glu Ala Arg
 275 280 285
 Arg Leu Arg Glu Leu Asn His Ile Arg Glu His Leu Asn Leu Pro Leu
 290 295 300
 Val Arg Ser Ala Ala Ala Glu Glu Pro Gly Ala Pro Leu Thr Thr Pro
 305 310 315 320
 Pro Val Leu Gln Gly Asn Gln Ala Arg Ser Ser Gly Tyr Phe Met Leu
 325 330 335
 Leu Ile Arg Ala Lys Leu Asp Ser Tyr Ser Ser Val Ala Thr Ser Glu
 340 345 350
 Gly Glu Ser Val Met Arg Glu His Ala Tyr Ser Arg Gly Arg Thr Arg
 355 360 365
 Asn Asn Tyr Gly Ser Thr Ile Glu Gly Leu Leu Asp Leu Pro Asp Asp
 370 375 380
 Asp Asp Ala Pro Ala Glu Ala Gly Leu Val Ala Pro Arg Met Ser Phe
 385 390 395 400
 Leu Ser Ala Gly Gln Arg Pro Arg Arg Leu Ser Thr Thr Ala Pro Ile
 405 410 415
 Thr Asp Val Ser Leu Gly Asp Glu Leu Arg Leu Asp Gly Glu Glu Val
 420 425 430
 Asp Met Thr Pro Ala Asp Ala Leu Asp Asp Phe Asp Leu Glu Met Leu
 435 440 445
 Gly Asp Val Glu Ser Pro Ser Pro Gly Met Thr His Asp Pro Val Ser
 450 455 460

Tyr Gly Ala Leu Asp Val Asp Asp Phe Glu Phe Glu Gln Met Phe Thr
 465 470 475 480

Asp Ala Met Gly Ile Asp Asp Phe Gly Gly
 485 490

<210> 206
 <211> 1443
 <212> DNA
 <213> HSV-2

<400> 206
 ccggccctg ggggccaaact acttctcctc catccgccag cccgtggtgc agcacgcccg 60
 cgagagcgcg gcgggggaga acgcgctgac ctacgcgctc atggcggggt acttcaagat 120
 gagecccggtg gccctgtatc accagctcaa gacgggcctc caccgccggg tcgggttcac 180
 cgtcgtgctg caggaccgct tcgtgaccga gaacgtgctg ttttccgagc gcgcgtcgga 240
 ggcgtacttt ctgggccagc tccaggtggc ccgccacgaa acgggcgggg gggtcagctt 300
 cacgctcacc cagccgcgcg gaaacgtgga cctgggtgtg ggctacaccg ccgtcgcggc 360
 cacggccacc gtccgcaacc ccgttacgga catgggcaac ctccccaaa acttttacct 420
 cgcccgcggg gcccccgcgc tgctagacaa cgccggccgc gtgtacctgc gcaacgcggg 480
 cgtggcgggg aaccggctgg ggccggccca gcccctcccg gtctttggct gcgccagggt 540
 gccgcggcgc gccggcatgg accacgggca ggatgcgctg tgtgagttca tcgccacccc 600
 cgtggccacg gacatcaact actttcgccg gccctgcaac ccgcggggac gcgcggccgg 660
 cggcgtgtac gcgggggaca aggaggggga cgtcatagcc ctcatgtacg accacggcca 720
 gagegacccg gcgcggccct tcgcggccac ggccaacccg tgggcgtcgc agcgggtctc 780
 gtacggggac ctgctgtaca acggggccta tcacctcaac ggggcctcgc ccgtcctcag 840
 cccctgcttc aagttcttca ccgcggccga catcacggcc aaacatcgct gcctggagcg 900
 tcttatcgct gaaacgggat cggcggtatc cacggccacc gctgccagcg acgtgcagtt 960
 taagcgcgcc ccgggggtgc gcgagctcgt ggaagaccgc tgcggcctgt ttcaggaagc 1020
 ctaccgcatc acctgcgcca gcgacccgc cctgctacgc agcgcgcgcg atggggaggc 1080
 ccacgcgcga gagaccact ttacgcagta tctcatctac gacgcctccc cgctaaaggg 1140
 cctgtctctg taaaattggc ttcttcttcc ccagctttcc ctccaccgcg ccgtcttttt 1200
 tttttttttt ttttttctgt ttggggcatt ggggttgatt ttccgacgtt gcttttacc 1260
 acacacaccc cctgtccccc cccccccggg ggggcttgga ctgggagccg cgattccgag 1320
 ggcaggtccc aataaaaccc agacccgagc tcggggggac tgattcttac ctggggctcc 1380
 tgcgcacgac agacctcccc gtgcgtgctg ctgagccctg ccccgcccc tctccacac 1440
 ggt 1443

<210> 207
 <211> 1313
 <212> DNA
 <213> HSV-2

<220>
 <221> misc_feature
 <222> 338
 <223> n = A,T,C or G

<400> 207
 gcgaacaat ccgtgatcat ggcgccgggc cgcgagacac cggaacgogg ggggtgcggga 60
 gggccggaaa gcgcaacgca accgggacga tgacgaaaca gagatggggg gcaccgaccg 120
 tgtgggagag ggggcggggc agggctcagc agcacgcacg gggaggtctg tcgtgcgcag 180
 gagccccagg taagaatcag tcccccgag ctccgggtctg ggttttattg ggacctgccc 240
 tcggaatcgc ggctccaggt ccaagccccc ccgggggggc ggggacaggg ggtgtgtgtg 300
 ggtaaaagca acgtcggaaa atcaaaccac atgccccnaa cagaaaaaaa aaaaaaaaaa 360

```

aaaaaagacg ggcgggtgga gggaaagctg gggaagaaga agccaatttt acagagacag 420
gccccttagc ggggagcggt cgtagatgag atactgcgta aagtgggtct ctgcgcgctg 480
ggcctcccca tcgcgggcgc tgcgtagcag ggcggggctc ctggcgaggg tgatcgggta 540
ggcttcctga aacaggccgc acgggtcttc cacgagctcg cggcaccocg gcgggcgctt 600
aaactgcacg tcgctggcag cggtgccgtt ggataccgcc gatcccgttt ccacgataag 660
acgctccagg cagcgatgtt tggccgtgat gtcggccgcg gtgaagaact tgaagcaggg 720
gctgaggacg ggcgagggcc cgttgaggtg ataggcccc ttgtacagca ggtccccgta 780
cgagaaccgc tgcgacgccc acgggttggc cgtggccgcg aaggggccgc cgggtcgct 840
ctggccgtgg tcgtacatga gggctatgac gtccccctcc ttgtcccccg cgtacacgcc 900
gccggccgcg cgtccccgcg ggttgaggg cggcgaaaag tagttgatgt ccgtggccac 960
gggggtggcg atgaactcac acacggcatc ctgcccgtgg tccatgccgg cgcgcgcg 1020
cacctgggcg cagccaaaga ccgggagggg ctggggccgc ccagccgggt ttcccgccac 1080
gaccgcgttg cgcaggtaca cggcgccgcg gttgtctagc agcggggggg cccgcgggcc 1140
gaggtaaaag ttttggggga ggttgcccat gtccgtaacg gggttgcgga cggtgccgt 1200
ggccgcgacg gcggtgtagc ccacaccag gtccacgttt ccgcgcggct gggtagcg 1260
gaagctgacc cccccgcgcg tttcgtggcg gggcacctgg agctggccca gaa 1313

```

<210> 208

<211> 1251

<212> DNA

<213> HSV-2

<400> 208

```

ggaggcgtcg tagatgagat actgcgtaaa gtgggtctct cgcgcgtggg cctccccatc 60
gcgggcgctg cgtagcaggg cggggctcgt ggcgcaggtg atcgggtagg ctctctgaaa 120
caggccgcac ggtcttcca cgagctcgcg gcaccccggc gggcgcttaa actgcacgtc 180
gctggcagcg gtggccgtgg ataccgccga tcccgtttcc acgataagac gctccaggca 240
gcgatgtttg gccgtgatgt cggccgcggt gaagaacttg aagcagggg tgaggacggg 300
cgaggccccg ttgaggtgat aggcccggt gtacagcagg tccccgtacg agaaccgctg 360
cgacgcccac gggttggccg tggccgcgaa gggccgcgcc gggtcgctct ggccgtggtc 420
gtacatgagg gctatgacgt cccctcctt gtcccccgcg tacacgccgc cggccgcgcg 480
tccccgcggg ttgcagggcc ggcgaaaagta gttgatgtcc gtggccacgg ggtggcgat 540
gaactcacac acggcatcct gcccggtgtc catgccggcg cgcgcgggca cctgggcgca 600
gccaaagacc gggaggggct gggccggccc cagccggttt ccgcgccaga ccgcgttgcg 660
caggtacacg gcggcccgct tgtctagcag cgggggggccc ccgcggccga ggtaaaagtt 720
ttgggtgagg ttgcccagt ccgtaacggg gttgcggacg gtggccgtgg ccgcgacggc 780
ggtgtagccc acaccaggt cagcttttc gcgcggctgg gtgagcgtga agctgacccc 840
cccgcccggt tcgtggcggg ccacctggag gcggtcctgc cgcacgacgg tgaaccggaa 960
ctcgaaaaac agcacgttct cggtcacgaa gcggtcctgc cgcacgacgg tgaaccggaa 960
cccggggtgg aggccgtct tgagctggtg atacagggcc acggggctca tcttgaagta 1020
ccccgccatg agcgcgtagg tcagcgcgtt ctcccccgcc gcgctctcgc gggcggtgctg 1080
caccacgggc tggcggtagg aggagaagta gttggcccc agggccgggg ggaccagggg 1140
gacgtggcgc gccaggtcgc gcagggccgg ggggaagttg ggcgcgttgg ccacgtggtc 1200
ggcggccgca aacagcgctt ggacgggcag gacgtagaag tattcgccat t 1251

```

<210> 209

<211> 881

<212> DNA

<213> HSV-2

<400> 209

```

ttctgggcca gctccaggtg gcccgccacg aaacgggccc ggggggtcagc ttcacgctca 60
cccagccgcg cggaaacgtg gacctgggtg tgggctacac cgccgtcgcg gccacggcca 120
ccgtccgcaa ccccgttacg gacatgggca acctcccca aaacttttac ctcgcccgcg 180
gggccccccc gctgctagac aacgcggccg ccgtgtacct gcgcaacgcg gtcgtggcgg 240
gaaaccggct ggggcccggc cagcccctcc cgggtctttg ctgcgccag gtgcgcggc 300
gcgcggcat ggaccacggg caggatgccg tgtgtgagtt catcgccacc cccgtggcca 360
cggacatcaa ctactttcgc cggccctgca acccgcgggg acgcgcggcc ggcggcgtgt 420

```

```

acgcggggga caaggagggg gacgtcatag ccctcatgta cgaccacggc cagagcgacc 480
cggcgcgggc cttcgcgggc acggccaacc cgtgggcgtc gcagcgggtc tcgtacgggg 540
acctgctgta caacgggggc tatcacctca acggggcctc gccgtcctc agcccctgct 600
tcaagtcttt caccgcgggc gacatcacgg ccaaacatcg ctgcctggag cgtcttatcg 660
tggaacggg atcggcggtg tccacggcca ccgctgccag cgacgtgcag tttaagcgcc 720
cgccggggtg ccgcgagctc gtggaagacc cgtgcggcct gtttcaggaa gcctaccgca 780
tcacctgcgc cagcgacccc gccctgctac gcagcgcccg cgatggggag gccacgcgc 840
gagagaccca ctttacgcag tatctcatct acgacgcctc c 881

```

<210> 210

<211> 4125

<212> DNA

<213> HSV-2

<400> 210

```

atggccgctc ctgcccgga cccccgggt taccggtaag ccgcggccat cctgcccacc 60
ggctccatcc tgagtacgat cgagggtggc tcccaccgca gactctttga ttttttcgcc 120
gccgtgcgct ccgacgaaaa cagcctgtat gacgtagagt ttgacgccct gctggggtcc 180
tactgcaaca ccctgtcgct cgtgcgcttt ctggagctcg gcctgtccgt ggctgcgtg 240
tgcaccaagt tcccgagct ggcttacatg aacgaagggc gtgtgcagtt cgaggccac 300
cagccccca tcgcccgga cggcccgac ccgctcgaag agcccgtgca taattacatg 360
acgaaggta tcgaccccg ggccctgaac gccgccttca gcctggccac cgaggccatt 420
gccctgctca cgggggaggc cctggacggg acgggcatta gcctgcacg ccagctgcgc 480
gccatccagc agctcgcgcg caacgtccag gccgtcctgg gggcggttga gcgcggcacg 540
gccgaccaga tgctgcacgt gctgttgag aaggcgccct ccctggccct gctgttgccc 600
atgcaacgat atctcgacaa cgggcgcctg gcgaccaggg ttgccgggc gaccctggtc 660
gccgagctga agcgagctt ttgcgacac agcttcttcc tgggcaaggc gggccatcgc 720
cgcgaggcca tcgaggcctg gctcgtggac ctgaccacgg cgacgcagcc gtccgtggcc 780
gtgccccgcc tgacgcacgc cgacacgcgc gggcgccgg tcgacggggt gctggtcacc 840
accgcccga tcaaacagcg cctcctgcag tccttctga aggtggagga caccgaggcc 900
gacgtgccgg tgacctacgg cgagatggtc ttgaacgggg ccaacctcgt cacggcgctg 960
gtgatgggca aggcctgcg gagcctggac gacgtgggcc gccacctgct ggatatgcag 1020
gaggagcaac tcgaggcgaa ccgggagacg ctggatgaac tcgaaagcgc ccccagaca 1080
acgcgcgtgc gcgcggatct ggtggccata ggcgacaggc tggctcttct ggaggccctg 1140
gagagacgca tctacgcgc caccaacgtg ccctaccccc tgggtgggcgc catggacctg 1200
acgttcgtcc tgcccctggg gctgttcaac ccggccatgg agcgtctcgc cgcgcacgcc 1260
ggggacctgg tgcccgcgcc cggccaccgg gagccccgc cgttccctcc ccggcagctg 1320
tttttttggg gaaaggacca caaggttctg cggctgtcca tggagaacgc ggtcggggc 1380
gtgtgtcatc cttcgtcat gaacatcgac ggcgcgtcg gggcggtgaa ccacgacccc 1440
gtcgaggccg cgaatcgtg cggggcgta gtcggggccc cggccggccc cggcgcgga 1500
atgcageagc gttttctgaa cgcctggcgg cagcgccctg cccacggccg ggtccgggtg 1560
gtcgccgagt gccagatgac cgcggagcag tcatgcagc ccgacaacgc caacctggct 1620
ctggagctgc accccgcgtt cgacttcttc gcggcgctgg ccgacgtcga gcttcccgcc 1680
ggcgaagtcc ccccgccgg tccgggggcg atccaggcca cctggcgctt ggtcaacggc 1740
aacctgcccc tggcgctgtg tccggtggcg ttctgtgacg cccggggcct ggagctcgcc 1800
gttggccgcc acgccatggc gccggctacc atagccggcg tccgcggggc gttcgaggac 1860
cgcagctacc cggcggtgtt ttacctgctg caagccgcga ttcacggcaa cgagcacgtg 1920
ttctgcgccc tggcgcggt cgtgactcag tgcatcacca gctactggaa caacacgca 1980
tgcgcgcgct tcgtgaacga ctactcgctg gtctcgta tctgaccta cctcgggggc 2040
gacctccccg aggagtgcac ggccgtgtat cgggacctgg tggccacgt cgaggccctg 2100
gccagctgg tggacgactt taccctgccg ggcccggagc tggcggggca ggctcaggcc 2160
gagctgaatc acctgatgc cgaccggcg ctgctgccgc ccctcgtgtg ggactgcgac 2220
ggccttatgc gacacgcggc cctggaccgc caccgagact gccggattga cgcggggggg 2280
cacgagccc tctacgcggc ggcgtgcaac gtggcgacgg ccgactttaa ccgcaacgac 2340
ggccggctgc tgcacaacac ccaggccgc gcggccgacg ccgcccagca ccggccgcac 2400
cggccggccg actggaccgt ccaccacaaa atctactatt acgtgctggt gccggccttc 2460
tcgcgggggc gctgctgcac cgcgggggtc cgttcgacc gcgtgtacgc cacgctgcag 2520
aacatggtgg tcccggagat cggccccggt gaggagtgc cgagcgatcc cgtgaccgac 2580

```



```

ccccccacc cgctgcatcc cgccaatctg gtggccaaca cgggtcaagcg catgttccac 2640
aacgggcgcg tcgtcgtcga cgggcccgcg atgctcacgc tgcaggtgct ggcgcacaac 2700
atggccgagc gcacgacggc gctgctgtgc tccgcggcgc ccgacgcggg cgccaacacc 2760
gcgtcgacgg ccaacatgcg catcttcgac ggggcgctgc acgccggcgt gctgctcatg 2820
gccccccagc acctggacca caccatccaa aatggogaat acttctacgt cctgcccgtc 2880
cacgcgctgt ttgcggggcg cgaccacgtg gccaacgcgc ccaacttccc cccggccctg 2940
cgcgacctgg cgcgcgacgt ccccttggtc ccccgggccc tggggggcaa ctactttccc 3000
tccatccgcc agcccggtgt gcagcacgcc cgcgagagcg cggcggggga gaacgcgctg 3060
acctacgcgc tcatggcggg gtacttcaag atgagccccg tggccctgta tcaccagctc 3120
aagacggggc tccaccccgg gttcgggttc accgtcgtgc ggcaggaccg ctctcgtgacc 3180
gagaacgtgc tgttttccga gcgcgcgtcg gaggcgtact ttctgggcca gctccagggtg 3240
gccccccacg aaacggggcg ggggttcaac ttcacgctca ccagccgcg cggaacgtg 3300
gacctgggtg tgggtacac cgccgtcgcg gccacgggca ccgtccgcaa ccccgttacg 3360
gacatgggca acctcccca aaacttttac ctgcggcgcg gggccccccc gctgctagac 3420
aacgcggccg ccgtgtacct gcgcaacgcg gtcgtggcgg gaaaccggct ggggccggcc 3480
cagccccctc cggctcttgg ctgcgcgcag gtgcgcggc gcgccgcat ggaccacggg 3540
caggatgccg tgtgtgagtt catcgccacc cccgtggcca cggacatcaa ctactttcgc 3600
cgccctgca acccgcgggg acgcgcggcc ggcggcgtgt acgcggggga caaggagggg 3660
gacgtcatag ccctcatgta cgaccacggc cagagcgacc cggcgcggcc cttcgcggcc 3720
acggccaacc cgtgggcgtc gcagcggttc tcgtacgggg acctgctgta caacggggcc 3780
tatcacctca acggggcctc gcccgtcctc agccctgct tcaagttctt caccgcggcc 3840
gacatcacgg ccaaacatcg ctgcctggag cgtctcatcg tgaaacggg atcggcggtg 3900
tccacggcca ccgtgccag cgacgtgcag ttttaagcgc cgccggggtg ccgcgagctc 3960
gtggaagacc cgtgcggcct gtttcaggaa gcctaccgca tcacctgcgc cagcgacccc 4020
gccctgctac gcagcgcccg cgatggggag gcccacgcgc gagagacca ctttacgcag 4080
tatctcatct acgacgcctc cccgctaaag ggctgtctc tgtaa 4125

```

<210> 211

<211> 293

<212> PRT

<213> HSV-2

<400> 211

```

Leu Gly Gln Leu Gln Val Ala Arg His Glu Thr Gly Gly Gly Val Ser
          5                      10                      15

```

```

Phe Thr Leu Thr Gln Pro Arg Gly Asn Val Asp Leu Gly Val Gly Tyr
          20                      25                      30

```

```

Thr Ala Val Ala Ala Thr Ala Thr Val Arg Asn Pro Val Thr Asp Met
          35                      40                      45

```

```

Gly Asn Leu Pro Gln Asn Phe Tyr Leu Gly Arg Gly Ala Pro Pro Leu
          50                      55                      60

```

```

Leu Asp Asn Ala Ala Ala Val Tyr Leu Arg Asn Ala Val Val Ala Gly
          65                      70                      75                      80

```

```

Asn Arg Leu Gly Pro Ala Gln Pro Leu Pro Val Phe Gly Cys Ala Gln
          85                      90                      95

```

```

Val Pro Arg Arg Ala Gly Met Asp His Gly Gln Asp Ala Val Cys Glu
          100                     105                     110

```

```

Phe Ile Ala Thr Pro Val Ala Thr Asp Ile Asn Tyr Phe Arg Arg Pro
          115                     120                     125

```

```

Cys Asn Pro Arg Gly Arg Ala Ala Gly Gly Val Tyr Ala Gly Asp Lys

```

130 135 140
 Glu Gly Asp Val Ile Ala Leu Met Tyr Asp His Gly Gln Ser Asp Pro
 145 150 155 160
 Ala Arg Pro Phe Ala Ala Thr Ala Asn Pro Trp Ala Ser Gln Arg Phe
 165 170 175
 Ser Tyr Gly Asp Leu Leu Tyr Asn Gly Ala Tyr His Leu Asn Gly Ala
 180 185 190
 Ser Pro Val Leu Ser Pro Cys Phe Lys Phe Phe Thr Ala Ala Asp Ile
 195 200 205
 Thr Ala Lys His Arg Cys Leu Glu Arg Leu Ile Val Glu Thr Gly Ser
 210 215 220
 Ala Val Ser Thr Ala Thr Ala Ala Ser Asp Val Gln Phe Lys Arg Pro
 225 230 235 240
 Pro Gly Cys Arg Glu Leu Val Glu Asp Pro Cys Gly Leu Phe Gln Glu
 245 250 255
 Ala Tyr Pro Ile Thr Cys Ala Ser Asp Pro Ala Leu Leu Arg Ser Ala
 260 265 270
 Arg Asp Gly Glu Ala His Ala Arg Glu Thr His Phe Thr Gln Tyr Leu
 275 280 285
 Ile Tyr Asp Ala Ser
 290

<210> 212
 <211> 1374
 <212> PRT
 <213> HSV-2

<400> 212
 Met Ala Ala Pro Ala Arg Asp Pro Pro Gly Tyr Arg Tyr Ala Ala Ala
 5 10 15
 Ile Leu Pro Thr Gly Ser Ile Leu Ser Thr Ile Glu Val Ala Ser His
 20 25 30
 Arg Arg Leu Phe Asp Phe Phe Ala Ala Val Arg Ser Asp Glu Asn Ser
 35 40 45
 Leu Tyr Asp Val Glu Phe Asp Ala Leu Leu Gly Ser Tyr Cys Asn Thr
 50 55 60
 Leu Ser Leu Val Arg Phe Leu Glu Leu Gly Leu Ser Val Ala Cys Val
 65 70 75 80
 Cys Thr Lys Phe Pro Glu Leu Ala Tyr Met Asn Glu Gly Arg Val Gln
 85 90 95
 Phe Glu Val His Gln Pro Leu Ile Ala Arg Asp Gly Pro His Pro Val

100	105	110
Glu Gln Pro Val His Asn Tyr Met Thr Lys Val Ile Asp Arg Arg Ala 115 120 125		
Leu Asn Ala Ala Phe Ser Leu Ala Thr Glu Ala Ile Ala Leu Leu Thr 130 135 140		
Gly Glu Ala Leu Asp Gly Thr Gly Ile Ser Leu His Arg Gln Leu Arg 145 150 155 160		
Ala Ile Gln Gln Leu Ala Arg Asn Val Gln Ala Val Leu Gly Ala Phe 165 170 175		
Glu Arg Gly Thr Ala Asp Gln Met Leu His Val Leu Leu Glu Lys Ala 180 185 190		
Pro Pro Leu Ala Leu Leu Leu Pro Met Gln Arg Tyr Leu Asp Asn Gly 195 200 205		
Arg Leu Ala Thr Arg Val Ala Arg Ala Thr Leu Val Ala Glu Leu Lys 210 215 220		
Arg Ser Phe Cys Asp Thr Ser Phe Phe Leu Gly Lys Ala Gly His Arg 225 230 235 240		
Arg Glu Ala Ile Glu Ala Trp Leu Val Asp Leu Thr Thr Ala Thr Gln 245 250 255		
Pro Ser Val Ala Val Pro Arg Leu Thr His Ala Asp Thr Arg Gly Arg 260 265 270		
Pro Val Asp Gly Val Leu Val Thr Thr Ala Ala Ile Lys Gln Arg Leu 275 280 285		
Leu Gln Ser Phe Leu Lys Val Glu Asp Thr Glu Ala Asp Val Pro Val 290 295 300		
Thr Tyr Gly Glu Met Val Leu Asn Gly Ala Asn Leu Val Thr Ala Leu 305 310 315 320		
Val Met Gly Lys Ala Val Arg Ser Leu Asp Asp Val Gly Arg His Leu 325 330 335		
Leu Asp Met Gln Glu Glu Gln Leu Glu Ala Asn Arg Glu Thr Leu Asp 340 345 350		
Glu Leu Glu Ser Ala Pro Gln Thr Thr Arg Val Arg Ala Asp Leu Val 355 360 365		
Ala Ile Gly Asp Arg Leu Val Phe Leu Glu Ala Leu Glu Arg Arg Ile 370 375 380		
Tyr Ala Ala Thr Asn Val Pro Tyr Pro Leu Val Gly Ala Met Asp Leu 385 390 395 400		
Thr Phe Val Leu Pro Leu Gly Leu Phe Asn Pro Ala Met Glu Arg Phe 405 410 415		

Ala Ala His Ala Gly Asp Leu Val Pro Ala Pro Gly His Pro Glu Pro
 420 425 430
 Arg Ala Phe Pro Pro Arg Gln Leu Phe Phe Trp Gly Lys Asp His Gln
 435 440 445
 Val Leu Arg Leu Ser Met Glu Asn Ala Val Gly Thr Val Cys His Pro
 450 455 460
 Ser Leu Met Asn Ile Asp Ala Ala Val Gly Gly Val Asn His Asp Pro
 465 470 475 480
 Val Glu Ala Ala Asn Pro Tyr Gly Ala Tyr Val Ala Ala Pro Ala Gly
 485 490 495
 Pro Gly Ala Asp Met Gln Gln Arg Phe Leu Asn Ala Trp Arg Gln Arg
 500 505 510
 Leu Ala His Gly Arg Val Arg Trp Val Ala Glu Cys Gln Met Thr Ala
 515 520 525
 Glu Gln Phe Met Gln Pro Asp Asn Ala Asn Leu Ala Leu Glu Leu His
 530 535 540
 Pro Ala Phe Asp Phe Phe Ala Gly Val Ala Asp Val Glu Leu Pro Gly
 545 550 555 560
 Gly Glu Val Pro Pro Ala Gly Pro Gly Ala Ile Gln Ala Thr Trp Arg
 565 570 575
 Val Val Asn Gly Asn Leu Pro Leu Ala Leu Cys Pro Val Ala Phe Arg
 580 585 590
 Asp Ala Arg Gly Leu Glu Leu Gly Val Gly Arg His Ala Met Ala Pro
 595 600 605
 Ala Thr Ile Ala Ala Val Arg Gly Ala Phe Glu Asp Arg Ser Tyr Pro
 610 615 620
 Ala Val Phe Tyr Leu Leu Gln Ala Ala Ile His Gly Asn Glu His Val
 625 630 635 640
 Phe Cys Ala Leu Ala Arg Leu Val Thr Gln Cys Ile Thr Ser Tyr Trp
 645 650 655
 Asn Asn Thr Arg Cys Ala Ala Phe Val Asn Asp Tyr Ser Leu Val Ser
 660 665 670
 Tyr Ile Val Thr Tyr Leu Gly Gly Asp Leu Pro Glu Glu Cys Met Ala
 675 680 685
 Val Tyr Arg Asp Leu Val Ala His Val Glu Ala Leu Ala Gln Leu Val
 690 695 700
 Asp Asp Phe Thr Leu Pro Gly Pro Glu Leu Gly Gly Gln Ala Gln Ala
 705 710 715 720

Glu Leu Asn His Leu Met Arg Asp Pro Ala Leu Leu Pro Pro Leu Val
 725 730 735
 Trp Asp Cys Asp Gly Leu Met Arg His Ala Ala Leu Asp Arg His Arg
 740 745 750
 Asp Cys Arg Ile Asp Ala Gly Gly His Glu Pro Val Tyr Ala Ala Ala
 755 760 765
 Cys Asn Val Ala Thr Ala Asp Phe Asn Arg Asn Asp Gly Arg Leu Leu
 770 775 780
 His Asn Thr Gln Ala Arg Ala Ala Asp Ala Ala Asp Asp Arg Pro His
 785 790 795 800
 Arg Pro Ala Asp Trp Thr Val His His Lys Ile Tyr Tyr Tyr Val Leu
 805 810 815
 Val Pro Ala Phe Ser Arg Gly Arg Cys Cys Thr Ala Gly Val Arg Phe
 820 825 830
 Asp Arg Val Tyr Ala Thr Leu Gln Asn Met Val Val Pro Glu Ile Ala
 835 840 845
 Pro Gly Glu Glu Cys Pro Ser Asp Pro Val Thr Asp Pro Ala His Pro
 850 855 860
 Leu His Pro Ala Asn Leu Val Ala Asn Thr Val Lys Arg Met Phe His
 865 870 875 880
 Asn Gly Arg Val Val Val Asp Gly Pro Ala Met Leu Thr Leu Gln Val
 885 890 895
 Leu Ala His Asn Met Ala Glu Arg Thr Thr Ala Leu Leu Cys Ser Ala
 900 905 910
 Ala Pro Asp Ala Gly Ala Asn Thr Ala Ser Thr Ala Asn Met Arg Ile
 915 920 925
 Phe Asp Gly Ala Leu His Ala Gly Val Leu Leu Met Ala Pro Gln His
 930 935 940
 Leu Asp His Thr Ile Gln Asn Gly Glu Tyr Phe Tyr Val Leu Pro Val
 945 950 955 960
 His Ala Leu Phe Ala Gly Ala Asp His Val Ala Asn Ala Pro Asn Phe
 965 970 975
 Pro Pro Ala Leu Arg Asp Leu Ala Arg Asp Val Pro Leu Val Pro Pro
 980 985 990
 Ala Leu Gly Ala Asn Tyr Phe Ser Ser Ile Arg Gln Pro Val Val Gln
 995 1000 1005
 His Ala Arg Glu Ser Ala Ala Gly Glu Asn Ala Leu Thr Tyr Ala Leu
 1010 1015 1020
 Met Ala Gly Tyr Phe Lys Met Ser Pro Val Ala Leu Tyr His Gln Leu

1025	1030	1035	1040
Lys Thr Gly Leu His Pro Gly Phe Gly Phe Thr Val Val Arg Gln Asp	1045	1050	1055
Arg Phe Val Thr Glu Asn Val Leu Phe Ser Glu Arg Ala Ser Glu Ala	1060	1065	1070
Tyr Phe Leu Gly Gln Leu Gln Val Ala Arg His Glu Thr Gly Gly Gly	1075	1080	1085
Val Asn Phe Thr Leu Thr Gln Pro Arg Gly Asn Val Asp Leu Gly Val	1090	1095	1100
Gly Tyr Thr Ala Val Ala Ala Thr Gly Thr Val Arg Asn Pro Val Thr	1105	1110	1115
Asp Met Gly Asn Leu Pro Gln Asn Phe Tyr Leu Gly Arg Gly Ala Pro	1125	1130	1135
Pro Leu Leu Asp Asn Ala Ala Ala Val Tyr Leu Arg Asn Ala Val Val	1140	1145	1150
Ala Gly Asn Arg Leu Gly Pro Ala Gln Pro Leu Pro Val Phe Gly Cys	1155	1160	1165
Ala Gln Val Pro Arg Arg Ala Gly Met Asp His Gly Gln Asp Ala Val	1170	1175	1180
Cys Glu Phe Ile Ala Thr Pro Val Ala Thr Asp Ile Asn Tyr Phe Arg	1185	1190	1195
Arg Pro Cys Asn Pro Arg Gly Arg Ala Ala Gly Gly Val Tyr Ala Gly	1205	1210	1215
Asp Lys Glu Gly Asp Val Ile Ala Leu Met Tyr Asp His Gly Gln Ser	1220	1225	1230
Asp Pro Ala Arg Pro Phe Ala Ala Thr Ala Asn Pro Trp Ala Ser Gln	1235	1240	1245
Arg Phe Ser Tyr Gly Asp Leu Leu Tyr Asn Gly Ala Tyr His Leu Asn	1250	1255	1260
Gly Ala Ser Pro Val Leu Ser Pro Cys Phe Lys Phe Phe Thr Ala Ala	1265	1270	1275
Asp Ile Thr Ala Lys His Arg Cys Leu Glu Arg Leu Ile Val Glu Thr	1285	1290	1295
Gly Ser Ala Val Ser Thr Ala Thr Ala Ala Ser Asp Val Gln Phe Lys	1300	1305	1310
Arg Pro Pro Gly Cys Arg Glu Leu Val Glu Asp Pro Cys Gly Leu Phe	1315	1320	1325
Gln Glu Ala Tyr Pro Ile Thr Cys Ala Ser Asp Pro Ala Leu Leu Arg	1330	1335	1340

Ser Ala Arg Asp Gly Glu Ala His Ala Arg Glu Thr His Phe Thr Gln
 1345 1350 1355 1360

Tyr Leu Ile Tyr Asp Ala Ser Pro Leu Lys Gly Leu Ser Leu
 1365 1370

<210> 213

<211> 1644

<212> DNA

<213> Herpes Simplex Virus Type 2

<400> 213

```

gctcgcgggg cggggttggt gttttttggt ggagtttggg tcgtatcgtg cctggcgcca 60
gcaccagaaa cgtcctggaa acgggtaacc tcgggcgagg acgtggtggt gcttccggcg 120
cccgcggggc cggaggaacg caccggggcc cacaaactac tgtgggcgcg ggaaccctg 180
gatgcctgcg gtccctgcg cccgtcgtgg gtggcgctgt ggcccccccg acgggtgctc 240
gagacggtcg tggatgcggc gtgcatgcgc gccccggaac cgctcgccat agcatacagt 300
ccccgttcc ccgcgggcga cgagggactg tttcggagt tggcgtggcg cgatcgcgta 360
gccgtggtea acgagagtct ggtcatctac ggggccctgg agacggacag cggctctgtac 420
accctgtccg tggtcggcct aagcgacgag gcgcgccaag tggcgtcggg ggttctggtc 480
gtggagcccg cccctgtgcc gaccccgacc cccgacgact acgacgaaga agacgacgcg 540
ggcgtgagcg aacgcacgcc ggtcagcgtt cccccccaa ccccccccg tcgtccccc 600
gtcgcgccccc cgacgcaccc tcgtgttatc cccgaggtgt cccacgtgcg cggggtaacg 660
gtccatatgg agaccccgga ggccattctg tttgcccccg gggagacggt tgggacgaac 720
gtctccatcc acgccattgc ccacgacgac ggtccgtacg ccatggacgt cgtctggatg 780
cggtttgacy tgcgcgtctc gtgcgcccag atgcggatct acgaagcttg tctgtatcac 840
ccgcagcttc cagagtgtct atctccggcc gacgcgccgt gcgcgctaag ttcttgggcy 900
taccgcctgg cggtcgcgag ctacgcgggc tgttccagga ctacgcccc gccgcgatgt 960
tttgccgagg ctgcgatgga accggtcccg gggttggcgt ggctggcctc caccgtcaat 1020
ctggaattcc agcacgcctc cccccagcac gccgcctct acctgtgcgt ggtgtacgtg 1080
gacgatcata tccacgcctg gggccacatg accatcagca ccgcggcgca gtaccggaac 1140
gcggtggtgg agcagcacct cccccagcgc cagcccgagc ccgtcgagcc caccgcgccg 1200
cacgtgagag cccccctcc ccgcgccctc gcgcgcggcc cgctgcgcct cggggcgatg 1260
ctgggggcyg ccctgttgct ggccgcctc gggctgtccg cgtgggcgtg catgacctgc 1320
tggcgaggcg gctcctggcg ggcggttaa agccgggct cggcgacggg cccacttac 1380
attcgcgtgg cggacgcga gctgtacgag gactggagtt cggacagcga gggggagcgc 1440
gacgggtccc tgtggcagga cctccggag agaccgcact ctccctccac aaatggatcc 1500
ggctttgaga tttatcacc aacggctccg tctgtatacc cccatagcga ggggcgtaaa 1560
tctcgcgcc cgctcaccac ctttggttcg ggaagcccg gccgtcgtca ctcccaggcc 1620
tcctatccgt ccgtcctctg gtaa 1644

```

<210> 214

<211> 1644

<212> DNA

<213> Herpes Simplex Virus Type 2

<400> 214

```

gctcgcgggg cggggttggt gttttttggt ggagtttggg tcgtatcgtg cctggcgcca 60
gcaccagaaa cgtcctggaa acgggtaacc tcgggcgagg acgtggtggt gcttccggcg 120
cccgcggggc cggaggaacg caccggggcc cacaaactac tgtgggcgcg ggaaccctg 180
gatgcctgcg gtccctgcg cccgtcgtgg gtggcgctgt ggcccccccg acgggtgctc 240
gagacggtcg tggatgcggc gtgcatgcgc gccccggaac cgctcgccat agcatacagt 300
ccccgttcc ccgcgggcga cgagggactg tttcggagt tggcgtggcg cgatcgcgta 360
gccgtggtea acgagagtct ggtcatctac ggggccctgg agacggacag cggctctgtac 420
accctgtccg tggtcggcct aagcgacgag gcgcgccaag tggcgtcggg ggttctggtc 480
gtggagcccg cccctgtgcc gaccccgacc cccgacgact acgacgaaga agacgacgcg 540

```

```

ggcgtgagcg aacgcacgcc ggtcagcggt ccccccgcaa ccccccccg tcgtcccccc 600
gtcgcccccc cgacgcaccc tcgtgttatt cccgaggtgt cccacgtgcg cggggtaacg 660
gtccatatgg agaccccgga ggcattctg tttgccccg gggagacgtt tgggacgaac 720
gtctccatcc acgccattgc ccacgacgac ggtccgtacg ccatggacgt cgtctggatg 780
cggtttgacg tgccgtcctc gtgcgcgag atgcggtatc acgaagcttg tctgtatcac 840
cgcagccttc cagagtgtct atctccggcc gacgcgcgt gcgcgtaag ttcctgggcg 900
taccgcctgg cggtcgcgag ctacgcgggc tgttcagga ctacgcccc gccgcgatgt 960
tttgccgagg ctgcgatgga accggtcccg gggttggcgt ggctggcctc caccgtcaat 1020
ctggaattcc agcagcctc ccccagcac gccggcctct acctgtgctg ggtgtacgcg 1080
gacgatcata tccacgcctg gggccacatg accatcagca ccgcggcgca gtaccggaac 1140
gcggtggtgg aacagcacct ccccagcgc cagcccgagc cgtcgagcc caccgcccg 1200
cacgtgagag ccccccattc cgcgcctcc gcgcgggcc cgtgcgctc cggggcggtg 1260
ctggggggcg cctgttgtgt ggccgcctc gggctgtccg cgtggcgctg catgacctgc 1320
tggcgaggc gctcctggcg ggcggttaa agccgggct cggcgacggg cccacttac 1380
attcgctgg cggacagcga gctgtacgag gactggagt cggacagcga gggggagcgc 1440
gacgggtccc tgtggcagga cctccggag agaccgact ctccctccac aaatggatcc 1500
ggctttgaga tcttatcacc aacggctccg tctgtatacc cccatagcga ggggcgtaaa 1560
tctcgccgcc cgtcaccac ctttggttcg ggaagcccg gccgtcgtca ctcccaggcc 1620
tcctatccgt ccgtcctctg gtaa 1644

```

<210> 215

<211> 547

<212> PRT

<213> Herpes Simplex Virus Type 2

<400> 215

```

Ala Arg Gly Ala Gly Leu Val Phe Phe Val Gly Val Trp Val Val Ser
      5                      10                      15

Cys Leu Ala Ala Ala Pro Arg Thr Ser Trp Lys Arg Val Thr Ser Gly
      20                      25                      30

Glu Asp Val Val Leu Leu Pro Ala Pro Ala Gly Pro Glu Glu Arg Thr
      35                      40                      45

Arg Ala His Lys Leu Leu Trp Ala Ala Glu Pro Leu Asp Ala Cys Gly
      50                      55                      60

Pro Leu Arg Pro Ser Trp Val Ala Leu Trp Pro Pro Arg Arg Val Leu
      65                      70                      75                      80

Glu Thr Val Val Asp Ala Ala Cys Met Arg Ala Pro Glu Pro Leu Ala
      85                      90                      95

Ile Ala Tyr Ser Pro Pro Phe Pro Ala Gly Asp Glu Gly Leu Tyr Ser
      100                     105                     110

Glu Leu Ala Trp Arg Asp Arg Val Ala Val Val Asn Glu Ser Leu Val
      115                     120                     125

Ile Tyr Gly Ala Leu Glu Thr Asp Ser Gly Leu Tyr Thr Leu Ser Val
      130                     135                     140

Val Gly Leu Ser Asp Glu Ala Arg Gln Val Ala Ser Val Val Leu Val
      145                     150                     155                     160

Val Glu Pro Ala Pro Val Pro Thr Pro Thr Pro Asp Asp Tyr Asp Glu
      165                     170                     175

```


Glu Asp Asp Ala Gly Val Ser Glu Arg Thr Pro Val Ser Val Pro Pro
 180 185 190
 Pro Thr Pro Pro Arg Arg Pro Pro Val Ala Pro Pro Thr His Pro Arg
 195 200 205
 Val Ile Pro Glu Val Ser His Val Arg Gly Val Thr Val His Met Glu
 210 215 220
 Thr Pro Glu Ala Ile Leu Phe Ala Pro Gly Glu Thr Phe Gly Thr Asn
 225 230 235 240
 Val Ser Ile His Ala Ile Ala His Asp Asp Gly Pro Tyr Ala Met Asp
 245 250 255
 Val Val Trp Met Arg Phe Asp Val Pro Ser Ser Cys Ala Glu Met Arg
 260 265 270
 Ile Tyr Glu Ala Cys Leu Tyr His Pro Gln Leu Pro Glu Cys Leu Ser
 275 280 285
 Pro Ala Asp Ala Pro Cys Ala Val Ser Ser Trp Ala Tyr Arg Leu Ala
 290 295 300
 Val Arg Ser Tyr Ala Gly Cys Ser Arg Thr Thr Pro Pro Pro Arg Cys
 305 310 315 320
 Phe Ala Glu Ala Arg Met Glu Pro Val Pro Gly Leu Ala Trp Leu Ala
 325 330 335
 Ser Thr Val Asn Leu Glu Phe Gln His Ala Ser Pro Gln His Ala Gly
 340 345 350
 Leu Tyr Leu Cys Val Val Tyr Val Asp Asp His Ile His Ala Trp Gly
 355 360 365
 His Met Thr Ile Ser Thr Ala Ala Gln Tyr Arg Asn Ala Val Val Glu
 370 375 380
 Gln His Leu Pro Gln Arg Gln Pro Glu Pro Val Glu Pro Thr Arg Pro
 385 390 395 400
 His Val Arg Ala Pro Pro Pro Ala Pro Ser Ala Arg Gly Pro Leu Arg
 405 410 415
 Leu Gly Ala Met Leu Gly Ala Ala Leu Leu Leu Ala Ala Leu Gly Leu
 420 425 430
 Ser Ala Trp Ala Cys Met Thr Cys Trp Arg Arg Arg Ser Trp Arg Ala
 435 440 445
 Val Lys Ser Arg Ala Ser Ala Thr Gly Pro Thr Tyr Ile Arg Val Ala
 450 455 460
 Asp Ser Glu Leu Tyr Ala Asp Trp Ser Ser Asp Ser Glu Gly Glu Arg
 465 470 475 480

Asp Gly Ser Leu Trp Gln Asp Pro Pro Glu Arg Pro Asp Ser Pro Ser
 485 490 495

Thr Asn Gly Ser Gly Phe Glu Ile Leu Ser Pro Thr Ala Pro Ser Val
 500 505 510

Tyr Pro His Ser Glu Gly Arg Lys Ser Arg Arg Pro Leu Thr Thr Phe
 515 520 525

Gly Ser Gly Ser Pro Gly Arg Arg His Ser Gln Ala Ser Tyr Pro Ser
 530 535 540

Val Leu Trp
 545

<210> 216

<211> 547

<212> PRT

<213> Herpes Simplex Virus Type 2

<400> 216

Ala Arg Gly Ala Gly Leu Val Phe Phe Val Gly Val Trp Val Val Ser
 5 10 15

Cys Leu Ala Ala Ala Pro Arg Thr Ser Trp Lys Arg Val Thr Ser Gly
 20 25 30

Glu Asp Val Val Leu Leu Pro Ala Pro Ala Gly Pro Glu Glu Arg Thr
 35 40 45

Arg Ala His Lys Leu Leu Trp Ala Ala Glu Pro Leu Asp Ala Cys Gly
 50 55 60

Pro Leu Arg Pro Ser Trp Val Ala Leu Trp Pro Pro Arg Arg Val Leu
 65 70 75 80

Glu Thr Val Val Asp Ala Ala Cys Met Arg Ala Pro Glu Pro Leu Ala
 85 90 95

Ile Ala Tyr Ser Pro Pro Phe Pro Ala Gly Asp Glu Gly Leu Tyr Ser
 100 105 110

Glu Leu Ala Trp Arg Asp Arg Val Ala Val Val Asn Glu Ser Leu Val
 115 120 125

Ile Tyr Gly Ala Leu Glu Thr Asp Ser Gly Leu Tyr Thr Leu Ser Val
 130 135 140

Val Gly Leu Ser Asp Glu Ala Arg Gln Val Ala Ser Val Val Leu Val
 145 150 155 160

Val Glu Pro Ala Pro Val Pro Thr Pro Thr Pro Asp Asp Tyr Asp Glu
 165 170 175

Glu Asp Asp Ala Gly Val Ser Glu Arg Thr Pro Val Ser Val Pro Pro
 180 185 190

Ala Thr Pro Pro Arg Arg Pro Pro Val Ala Pro Pro Thr His Pro Arg
 195 200 205
 Val Ile Pro Glu Val Ser His Val Arg Gly Val Thr Val His Met Glu
 210 215 220
 Thr Pro Glu Ala Ile Leu Phe Ala Pro Gly Glu Thr Phe Gly Thr Asn
 225 230 235 240
 Val Ser Ile His Ala Ile Ala His Asp Asp Gly Pro Tyr Ala Met Asp
 245 250 255
 Val Val Trp Met Arg Phe Asp Val Pro Ser Ser Cys Ala Glu Met Arg
 260 265 270
 Ile Tyr Glu Ala Cys Leu Tyr His Pro Gln Leu Pro Glu Cys Leu Ser
 275 280 285
 Pro Ala Asp Ala Pro Cys Ala Val Ser Ser Trp Ala Tyr Arg Leu Ala
 290 295 300
 Val Arg Ser Tyr Ala Gly Cys Ser Arg Thr Thr Pro Pro Pro Arg Cys
 305 310 315 320
 Phe Ala Glu Ala Arg Met Glu Pro Val Pro Gly Leu Ala Trp Leu Ala
 325 330 335
 Ser Thr Val Asn Leu Glu Phe Gln His Ala Ser Pro Gln His Ala Gly
 340 345 350
 Leu Tyr Leu Cys Val Val Tyr Ala Asp Asp His Ile His Ala Trp Gly
 355 360 365
 His Met Thr Ile Ser Thr Ala Ala Gln Tyr Arg Asn Ala Val Val Glu
 370 375 380
 Gln His Leu Pro Gln Arg Gln Pro Glu Pro Val Glu Pro Thr Arg Pro
 385 390 395 400
 His Val Arg Ala Pro His Pro Ala Pro Ser Ala Arg Gly Pro Leu Arg
 405 410 415
 Leu Gly Ala Val Leu Gly Ala Ala Leu Leu Leu Ala Ala Leu Gly Leu
 420 425 430
 Ser Ala Trp Ala Cys Met Thr Cys Trp Arg Arg Arg Ser Trp Arg Ala
 435 440 445
 Val Lys Ser Arg Ala Ser Ala Thr Gly Pro Thr Tyr Ile Arg Val Ala
 450 455 460
 Asp Ser Glu Leu Tyr Ala Asp Trp Ser Ser Asp Ser Glu Gly Glu Arg
 465 470 475 480
 Asp Gly Ser Leu Trp Gln Asp Pro Pro Glu Arg Pro Asp Ser Pro Ser
 485 490 495
 Thr Asn Gly Ser Gly Phe Glu Ile Leu Ser Pro Thr Ala Pro Ser Val

500 505 510

Tyr Pro His Ser Glu Gly Arg Lys Ser Arg Arg Pro Leu Thr Thr Phe
515 520 525

Gly Ser Gly Ser Pro Gly Arg Arg His Ser Gln Ala Ser Tyr Pro Ser
530 535 540

Val Leu Trp
545

<210> 217
<211> 1376
<212> DNA
<213> Herpes simplex virus

<400> 217

cgtgcaagtg	aagcacatcg	accgtgttgt	gtccccgagc	gtctcttcg	cccccccgcc	60
gagcgcccc	gacgogagtc	tgccgcccc	ggggctccag	gaggccgccc	cgccgggccc	120
cccgtcagg	gagctgtggt	gggtgttcta	cgccggcgac	cgggcgctgg	aggagcccca	180
cgccgagtcg	ggattgacgc	gcgaggaggt	ccgcgcctg	catgggttcc	gggagcaggc	240
gtggaagctg	tttgggtcgg	tgggggctcc	gcgggcgttt	ctcggggccg	cgctggccct	300
gagcccgaac	caaaagctcg	ccgtctacta	ctatctcacc	caccgggagc	ggcgcatgtc	360
ccccctcccc	gcgctcgtgc	ggctcgtcgg	tccgtacatc	cagcgccacg	gcctgtacgt	420
tcccgcgccc	gacgaaccga	cgttggccga	tgccatgaac	gggctgttcc	gcgacgcgct	480
ggcgcccggg	accgtggccg	agcagtcct	catgttcgac	ctcctcccgc	ccaaggacgt	540
gccggtgggg	agcgacgcgc	gggcccagag	cgccgccttg	ctgcgctttg	tggactcgca	600
acgcctgacc	ccgggggggt	ccgtctcgcc	cgagcacgtc	atgtacctcg	gcgcgttcct	660
ggcgctgttg	tacgcccggc	acggacgcct	ggccgcggcc	acgcataccg	cgcgcccgac	720
ggcgctgacg	tccctggtec	tgaccgtggg	ggacgtcgac	cggatgtccg	cgtttgaccg	780
cgggccggcg	ggggcggtcg	gccgcacgcg	aaccgccggg	tacctggacg	cgctgcttac	840
cgtttgccctg	gctcgcgccc	agcacggcca	gtctgtgtga	gatatcccaa	taaagtgcag	900
tcgttttcta	accacggat	gccgttgtat	gcctatacgg	gggactatgg	ggggggggga	960
aaggaaagga	aacaggaatg	gagaaggga	aggaacagag	gcggtagcgg	acgcacggcg	1020
gacacaataa	caaacagacc	gcggacacgg	agggagtcgg	ttgggttggg	cgtggagccc	1080
gctgcgtcca	cacacccgtt	tattcgcgtc	tccacaaaaa	tgggacgcac	gttcggacca	1140
ccctgaggat	gcccgcaggg	gccgcggtga	tcataacgac	ccccagcgcg	gacgcggcca	1200
gaaacccggg	ggcgatggtg	gcgatgggca	gcgtgtcaaa	ggccagcaga	tgaatcacag	1260
ttccgttggg	gaacaacaac	agggccacgg	acggcacgtc	gctggaaaac	acgttcgggg	1320
tgcccgcac	cggcccctgg	gccagctgct	gttggttgge	atccgtgtcc	accage	1376

<210> 218
<211> 1599
<212> DNA
<213> Herpes simplex virus

<400> 218

atggagctca	gctatgccac	caccctgcac	caccgggacg	ttgtgtttta	cgtcacggca	60
gacagaaacc	gcgcctactt	tgtgtgcggg	gggtccgttt	attccgtagg	gcggcctcgg	120
gattctcagc	cgggggaaat	tgccaagttt	ggcctggtgg	tccgggggac	aggccccaaa	180
gaccgcatgg	tcgccaacta	cgtacgaagc	gagctccgcc	agcgcgccct	gcgggacgtg	240
cgggccgctg	gggaggacga	ggtgttcctg	gacagcgtgt	gtctgctaaa	cccgaacgtg	300
agctccgagc	gagacgtgat	taataccaac	gacgttgaag	tgttggaaga	atgcctggcc	360
gaatactgca	cctcgtgcg	aaccagcccg	ggggtgctgg	tgaccggggg	gcgcgtgcgc	420
gcgcgagaca	gggtcatcga	gctattttgag	caccggcgca	tcgtcaacat	ttcctcgcgc	480

ttcgcgtaca	ccccctcccc	ctacgtattc	gccttggccc	aggcgcacct	cccccggtc	540
ccgagctcgc	tggagccct	ggtgagcggc	ctgtttgacg	gcattcccgc	cccgcgccag	600
cccctggacg	cccgcgaccg	gcgcacggat	gtcgtgatca	cgggcacccg	cgccccaga	660
ccgatggccg	ggaccggggc	cgggggcgcg	ggggccaagc	gggccaccgt	cagcgagttc	720
gtgcaagtga	agcacatcga	ccgtgttggt	tccccgagcg	tctcttcgcg	ccccccgccc	780
agcgcccccg	acgcgagttc	gccgcccccg	gggtctccag	aggccgcccc	gccgggcccc	840
ccgctcaggg	agctgtgggt	ggtgtttctac	gccggcgacc	gggcgctgga	ggagccccac	900
gcogagtcgg	gattgacgcg	cgaggagggtc	cgcgcctgtc	atgggttcgg	ggagcaggcg	960
tggagagctgt	ttgggtcggg	gggggctccg	cgggcgtttc	tcggggccgc	gctggccctg	1020
agcccagacc	aaaagctcgc	cgtctactac	tatctcatcc	accgggagcg	gcgcatgtcc	1080
cccttccccg	cgctcgtgcg	gctcgtcggg	cgggtacatcc	agcgccacgg	cctgtacggt	1140
cccgcccccg	acgaaccgac	gttgcccgat	gccatgaacg	ggctgttcgg	cgacgcgctg	1200
gcggccggga	ccgtggccga	gcagctcctc	atgttcgacc	tcctcccggc	caaggacgtg	1260
ccggtgggga	gcgacgcgcg	ggccgacagc	gccgcctgtc	tgcgctttgt	ggactcgcaa	1320
cgcctgaccc	cgggggggtc	cgtctcgcgc	gagcacgtca	tgtacctcgg	cgcgttccctg	1380
ggcgtgttgt	acgcgggcca	cggacgcctg	gccgcggcca	cgcataccgc	gcgcctgacg	1440
ggcgtgacgt	ccctggtcct	gaccgtgggg	gacgtcgacc	ggatgtccgc	gtttgaccgc	1500
gggcccggcg	gggcggctgg	ccgcaecgga	accgcggggt	acctggacgc	gctgcttacc	1560
gtttgcctgg	ctcgcgcccc	gcacggccag	tctgtgtga			1599

<210> 219

<211> 746

<212> DNA

<213> Herpes simplex virus

<400> 219

cgctcgcgcc	ctcatcgctc	gcctgaccgt	ggagagcggg	tttgaggcca	acgtggccgc	60
ggtcgtcggg	tcccgcacga	cgggccttgg	ggggacgcgc	gtgtccctca	agcttatgcc	120
ctcgcactac	agcccgctcg	tctacgtctt	tcacggcggc	cgccacctgg	ccccagcac	180
ccaggccccc	aacctgacgc	ggctctgcga	gcgggcgcgc	cgccattttg	gcttctcgga	240
ctacgcgcgc	cggccctgcg	acctaaagca	cgagacgacg	ggggacgcgc	tgtgcgagcg	300
gctcggactg	gacccggacc	gggccttgct	gtatctggtg	atcacggagg	gcttccggga	360
ggcgtgtgct	atcagcaaca	ccttcctgca	cctgggcggg	atggacaagg	taaccatcgg	420
ggacgcggag	gttcaccgca	tcccgtgta	tccgctgcag	atgtttatgc	cggattttag	480
ccgggtcatc	gccgatccgt	ttaactgcaa	ccaccgatcg	atcggggaga	atttcaacta	540
cccccttcgg	tttttcaacc	gccccctcgc	cgcctgctg	ttcgaggcgg	tcgtcgggcc	600
cgcgcgcgtg	gcctgcgctg	ctcggaacgt	ggacgcctg	gcccgcgcgg	ccgcccacct	660
ggcgttcgac	gaaaaccacg	agggcgcggc	cctccccgcc	gacattacgt	tcacggcctt	720
cgaggccagc	cagggtaagc	cccagc				746

<210> 220

<211> 823

<212> DNA

<213> Herpes simplex virus

<400> 220

ggcctggagc	tgctgtcgtc	actgtcggcg	cgcagcggcg	acgcggacgt	cgcctgcgcg	60
cccctcatcg	tcggcctgac	cgtggagagc	gggtttgagg	ccaacgtggc	cgcggtcgtc	120
gggtcccga	cgacgggcct	tggggggacc	gcggtgtccc	tcaagcttat	gcctcgcac	180
tacagcccgt	ccgtctacgt	ctttcacggc	ggccgccacc	tggccccag	caccaggcc	240
cccaacctga	cgcggtctcg	cgagcggggc	cgccgccatt	ttggcttctc	ggactacgcg	300
ccccggccct	g'cgacctaaa	gcacgagacg	acgggggacg	cgtgtgtcga	gcggtcggga	360
ctggaccggg	accgggcctt	gctgtatctg	gtgatcacgg	agggcttccg	ggaggccgtg	420
tgcatcagca	acaaccttct	gcacctgggc	gggatggaca	aggtaaccat	cggggacgcg	480
gaggttcacc	gcatecccggt	gtatccgctg	cagatgttta	tgccggattt	tagccgggtc	540
atcgccgatc	cgtttaactg	caaccaccga	tcgatcgggg	agaatttcaa	ctacccctt	600
ccgtttttca	accgccccct	cgcgcgcctg	ctgttcgagg	cggctcgtcg	gcccgcgcgc	660
gtggccctgc	gtgctcggaa	cgtggacgcc	gtggcccgcg	cggccgcccc	cctggcgctt	720

gacgaaaacc	acgagggcgc	ggccctcccc	gccgacatta	cgttcacggc	cttcgagggc	780
agccagggtg	agccccagcg	ggcgcgcgcg	gacgcccggg	acc		823

<210> 221

<211> 3591

<212> DNA

<213> Herpes simplex virus

<400> 221

atggacacca	agcccaaaac	gacgaccacc	gtcaagggtc	ctccggggcc	gatgggggtac	60
gtgtatggcc	gcgcctgtcc	ggccgaaggc	ctggagctgc	tgtcgtact	gtcggcgcg	120
agcggcgacg	cggacgtcgc	cgctcgcccc	ctcatcgtcg	gcctgaccgt	ggagagcggg	180
tttgaggcca	acgtggccgc	ggctcgtcgg	tcccgacaga	cgggccttgg	ggggaccgcg	240
gtgtccctca	agcttatgcc	ctcgactac	agcccgctcg	tttacgtctt	tcacggcgcc	300
cgcacactgg	ccccagcac	ccaggcccc	aacctgacgc	ggctctgcga	gcgggcgcgc	360
ccgcattttg	gcttcgcgga	ctacgcgccc	cggccctgcg	acctaaagca	cgagacgacg	420
ggggacgcgc	tgtgcgagcg	gctcggactg	gaccgcggacc	gggccttgcg	gtatctgggtg	480
atcacggagg	gcttcgggga	ggcgtgtgc	atcagcaaca	ccttcctgca	tctgggcggg	540
atggacaagg	taaccatcgg	ggacgcggag	gttcaccgca	tcccgctgta	tccgctgcag	600
atgtttatgc	cggattttag	ccgggtcatc	gccgatccgt	ttaactgcaa	ccaccgatcg	660
atcggggaga	atttcaacta	cccccttcgc	tttttcaacc	gccccctcgc	ccgcctgctg	720
ttcgaggcgg	tcgtcggggc	cgcgcgcgtg	gccctgcgtg	cccggaacgt	ggacgcgcgtg	780
gcccgcgcgg	ccgccacact	ggcgttcgac	gaaaaccacg	agggcgcggc	cctccccgcc	840
gacattacgt	tcacggcctt	cgaggccagc	cagggttaagc	cccagcgggg	cgcgcgcgac	900
gccgggaaca	agggcccggc	ggcggggttt	gagcagcgcc	tggcctcggg	gatggcgggc	960
gacgcgcgcc	tggccctgga	gtccatcgtg	tcgatggcgg	tctttgacga	gccgcccccc	1020
gacatcacca	cgtggccctt	gcttgagggc	caggaaaacg	ccgcggcccg	cgcggcgcc	1080
gtcggggcgt	acctggcgcg	cgcgcggggc	ctcgtgggcg	ccatggtgtt	cagcaccaat	1140
tcggccctcc	acctgaccga	ggtggacgac	gccgggcccg	cggaccccaa	ggaccacagc	1200
aagccctcct	tttaccgctt	ctttctcgtg	cccggcacgc	acgtggcggc	caaccgcgag	1260
ctggaccgcg	agggccacgt	ggtgcccggg	tacgaggggc	gccccacggc	gcccctcgtc	1320
ggcggaaacc	aggagtccgc	cggcgagcac	ctggccatgc	tgtgcggggt	ctctccggcc	1380
ctgctggcca	agatgctgtt	ttacctggag	cgtcgcgacg	gcggagtcac	cgctcggcg	1440
caggagatgg	acgtgtttcg	ctacgtcgcg	gactcggggc	agaccgacgt	gccctgcaac	1500
ctgtgtacct	tcgagaccgc	ccacgcctgc	gcgcacacga	cgtgatgcg	gctgcggggc	1560
cgccacccca	agttcgcgag	cgcgcggcgg	ggggccattg	gggtcttcgg	aacgatgaac	1620
agcgcgtaca	gcgactgcga	cgtgctgggg	aactacgccc	ccttctcggc	cctgaagcgc	1680
gcggacgggt	ccgagaacac	ccggaccatc	atgcagagga	cgtaccgcgc	ggccaccgag	1740
cgcgtcatgg	ccgagctcga	ggccctgcag	tacgtggacc	agggcgtgcc	cacggccctg	1800
gggaggctgg	agacgattat	cggcaaccgc	gaggccctgc	acacggtggt	gaacaacatc	1860
aagcagctgg	tggatcgcga	ggtggaacag	ctgatgcgca	acctgatcga	ggggcgcaac	1920
tttaagttcc	gcgacgggtc	ggccgaggcc	aaccacgcca	tgtcgtctgc	gctggacccg	1980
tacacctgcg	ggccgtgccc	cctgctgcag	ctcctcgcgc	ggcgggtccaa	cctcgccgtg	2040
taccaggacc	tggccctgag	ccagtgccac	ggggtgttcg	ccgggcagtc	ggtcgagggg	2100
cgcaactttc	gcaatcaatt	ccaaccggtg	ctgcgcgcgc	gcgtcatgga	cctgttcaac	2160
aacgggttcc	tgtcggcgaa	gacgtgacg	gtcgcgtctc	cggagggggc	ggctatctgc	2220
gccccagcc	taacggccgg	ccagacggcc	cccgcggaga	gcagcttcga	gggcgacgtt	2280
gcccgcgtga	cccttgggtt	tcccaaggaa	ctgcgcgtca	agagccgcgt	gttgttcgcg	2340
ggcgcgagcg	ccaacgcgtc	cagggccgcc	aaggcgcggg	tcgccagcct	ccagagcgcc	2400
taccagaagc	ccgacaagcg	cgtggacatc	ctcctcggac	cgtgggctt	cctgctgaag	2460
cagttccacg	cgtcatctt	tcccaacggc	aagccccggg	ggtcgaacca	gccgaaccgc	2520
cagtgggtct	ggaccgccct	ccagcgcaac	cagctccccg	cccggctctt	gtcgcgcgag	2580
gacatagaga	ccatcgcgtt	cattaagagg	ttctccctgg	actacggcgc	gattaacttt	2640
atcaacctgg	cccccaataa	cgtgagcgaa	ctggccatgt	actacatggc	caaccagatc	2700
ttgaggtaact	gcgaccactc	gacgtacttc	ataaacaccc	tcacggccgt	catcgccggg	2760
tcccgcgcgt	cccccagcgt	gcaggcggcg	gccgcctggg	ccccgcaggg	cggcgcgggc	2820
ctggaggccg	gggcccgcgc	gctgatggac	tccctcgacg	cgcaccccg	cgcctggacg	2880
tccatgttcg	cagactgcaa	cctgctgcgg	cccgtgatgg	ccgcgcgcc	gatggctcgtg	2940

ctgggcctca	gcatcagcaa	atactacggc	atggcgggca	acgaccgcgt	gtttcaggcc	3000
gggaactggg	ccagcctgct	gggcggcaaa	aacgcgtgcc	cgctgctgat	ctttgaccgc	3060
acccgcaagt	ttgtcctggc	ctgccccagg	gccgggtttg	tgtgcgcggc	gtcagatctc	3120
gggggcgggg	cccacgagca	ttcgctgtgc	gagcagctcc	ggggcatcat	tgcggagggc	3180
ggggcgggcg	tcgccagcag	cgtgtttgtg	gcgaccgtca	agagcctggg	ccccgcacc	3240
cagcagctgc	agatcgagga	ctggctggcg	ctgctggagg	acgagtacct	gagcgaggaa	3300
atgatggagt	ttaccaccgc	cgccctggag	cgccggcgatg	gcgagtggtc	gaccgacgcg	3360
gccctggagg	tggccacga	ggcggaggcc	ctggtcagcc	agctcggcgc	cgccggggag	3420
gtgtttaact	ttggggattt	cgccgacgag	gatgaccacg	cgccgtcggt	cgccgggttg	3480
gcgccggcg	cgccggcg	cgccgtcgcc	cgaaagcggg	cgtttcaagg	cgacgatccg	3540
tttggggagg	gaccaccgga	gaaaaaggac	ctgacgttgg	atatgtctta	g	3591

<210> 222

<211> 761

<212> DNA

<213> Herpes simplex virus

<400> 222

gccccgcctc	gccccgcctc	gccccgcctc	gccccgcctc	gccccgcctc	gccccgcctc	60
gccagacgcg	gagggcgaac	cggaccagat	ggaaaacacg	tatctgctgc	ccgacgatga	120
cgccgccatg	cccgccggcg	tccggcttgg	cgccaccccc	gccccgcgaca	ccaccgcccgc	180
cgccctggccg	gccccgaagcc	acgccccccg	cgccccctcc	gagggacgcag	attccattta	240
cgagtcgggtg	agcgaggatg	gggggcgcgt	ctacgaggag	atccccctggg	ttcgggtata	300
cgaaaacatc	tgccctcgcc	ggcaagacgc	cgccggggcc	gccccgcggg	gagacgcccc	360
ggactccccg	tacatcgagg	cggaataatcc	cctgtacgac	tggggcggtg	ctgccccctt	420
ctccccctcg	ggggccacac	gcgccccgga	cccgggacta	agcctgtcgc	ccatgcccgc	480
ccgccccccg	accaacgcgc	tggccaacga	cgccccgaca	aacgtgcgcg	ccctcagcgc	540
cctgttgacg	aagctcaaac	gccccgcgaca	ccagagccat	taaaaaaatg	cgaccgcccg	600
ccccaccgtc	tcggtttccg	gcccccttcc	ccgtatgtct	gtttttcaata	aaaagtaaca	660
aacagagaaa	aaaaaacagc	gagttccgca	tgggtttgtc	tacgcaatta	gctgtttatt	720
gttttttttt	tggggggggg	aagagaaaaa	gaaaaaagga	g		761

<210> 223

<211> 1031

<212> DNA

<213> Herpes simplex virus

<400> 223

cggaccgcggc	gggacaacgc	acccccgcgc	ctccccgcggg	ctcgccccca	ctcgaccccc	60
gccccccctc	ggaggtgcag	gcgccaccgc	gaggacctcc	ccgagccccc	gcacgtcgac	120
gccccgcgac	gggggtccga	gccctgcgcc	ggccggccgg	ccacgtatta	cacgcataatg	180
gccccggcg	ccccgcgcct	cccgccccgc	aacccccgcg	cccccgagca	gcggccggca	240
gccccgcgcg	gccccgcctgc	ggctcagcgc	gagggccgcg	gggtctacga	cgccgtgcgg	300
acctgggggc	cggacgcgga	ggccgaaccg	gaccagatgg	aaaacacgta	tctgctgccc	360
gacgatgacg	ccgccatgcc	cgccggcgctc	ggccttggcg	ccacccccgc	cgccgacacc	420
accgcccgcg	ccgcctggcc	ggccgaaagc	cacgcccccc	gcgccccctc	cgaggacgca	480
gattccattt	acgagtcggg	ggcgaggatg	ggggggcgcg	tctacgagga	gatccccctgg	540
gttcgggtat	acgaaaacat	ctgccctcgc	cgccgtcttg	ccggcggggc	cgccctgccc	600
ggagacgccc	cggactcccc	gtacatcgag	gcgggaaatc	ccctgtacga	ctggggcggg	660
tctgccctct	tctccccctc	gcgggccaca	cgccccccgc	acccgggact	aagcctgtcg	720
cccatgccc	cccgccccgc	gaccaacgcg	ctggccaacg	acggccccgc	gaacgtcgcc	780
gccctcagcg	ccctgttgac	gaagctcaaa	cgccggccgac	accagagcca	ttaaaaaaat	840
gcgaccgccc	gccccaccgt	ctcggtttcc	ggcccccttc	cccgtatgtc	tgtttcaata	900
aaaagtaaca	aacagagaaa	aaaaacagcg	agttccgcgt	ggtttgcgt	acgcaattag	960
ctgtttattg	ttttttttgg	ggggggggga	agagaaaaag	aaaaaaggag	ggtattggcc	1020
aaacgcccgc	g					1031

<210> 224

<211> 2169

<212> DNA

<213> Herpes simplex virus

<400> 224

```

atgcaacgcc gggcgcgccg cgcgagctcc ctgcggtgctg cgcggtgcct gacgccggcc      60
aacctcatcc gggcgcccaa cgcgggcgct cccgagcggc gcattcttcg cggtgtctg      120
ctccccaccg cgagggggt cctcagcgcg gccgtggcg tectgcggca gcgcgccgac      180
gacctgcagc cggtgtttct gaccggcgcc gatcgagcg tccggctggc ggcgcgccac      240
cataacaccg tccccgagag cctgatcgta gacgggctcg ccagcgaccc gcaactacgac      300
tacatccggc actacgcgtc ggccgccaag caggcgctcg gcgaggtgga gctgtcgggc      360
ggccagctga gccgcgccat cctagcgag tactggaagt acctccagac ggtcgtgccc      420
agcggtctgg acatccccga cgaccggcg ggcgactgag acccagcct gcacgtgctg      480
ctgcggccca cctgtctccc gaagctgctg gtgcgcgccc cgttcaagag cggggccgccc      540
gcggccaagt acgcgcgcgc ggtggcgggg ttgcgcgacg cgcccacag gctccagcag      600
tacatgttct ttatgcgccc cgcagaccgc agccggcgga gcacggacac cgactgcgg      660
ctgagcgagc tcttgcccta cgtctccgtg ttgtaccatt gggcctcgtg gatgctgtgg      720
acggcgagca agtacgtgtg tcgcgcctg ggccccgcgc atcgccggtt cgtggcgctc      780
agcgggagtc tggaggcgcc cgcgagagc tttgcgcgcc acctggaccg cggggccagc      840
ggcaccacgg gctcgatgca gtgcattggc ctgcggcgcg cgtcagcga cgtcctgggc      900
cacctgacgc gcctggccca cctgtgggag accggcaagc gcagcgcgcg cactacggg      960
atcgtggacg ccattcgttc gaccgtcgag gttctatcca tagtccacca ccacgccag      1020
tatataatta acgcgacgct taccgggtat gtgctctggg cctccgacag cctgaacaac      1080
gagtacctta cggcgcggtt ggacagccag gaggccttct gcaggaccgc cgccccctg      1140
ttccccacga tgaccgcccc gagctgggcc cggatggaac tcagcatcaa gtctgtgttc      1200
ggggccgccc tggccccgga cctgcttcgg agcggaaacc cgtcgcccca ctacgagtc      1260
atcctgcgcc tcgcgcgctc cgccccaccg gggggcgcg gcgcggtcgg cgggagctgc      1320
cgggacaaga tacaacggac ccggcgcgac aacgcacccc cgccgctccc ccgggctcgc      1380
ccccactcga ccccgcggc ccctcgagg tgcaggcgcc acccgagga cctccccgag      1440
ccccgcacg tcgacgcggc cgaccgggt cccgagccct gcgcggcgcc gccggccacg      1500
tattacacgc atatggccgg ggcgcccccg cgcctccgc cccgcaacc cgcgcccccc      1560
gagcagcggc cggcagccgc ggcgcgccc ctgcggtctc agcgcgaggc cgccggggtc      1620
tacgacgcgg tgcggacctg ggggcccggc gcggaggccg aaccggacca gatggaaaac      1680
acgtatctgc tgcccgacga tgacgcgcgc atgcccgcgg gcgtcggtt tggcgccacc      1740
cccgccgcgc acaccaccgc cgccgcgcgc tggccggccg aaagccacgc ccccgcgcc      1800
cctccgagg acgcagattc catttacgag tcggtggcg aggatgggg gcgcgtctac      1860
gaggagatcc cctgggttcg ggtatacga aacatctgcc ctgcgcggc tcttgccggc      1920
ggggccgccc tgcggggaga cgccccggac tccccgtaca tcgaggcgga aaatccccctg      1980
tacgactggg gcgggtctgc cctcttctcc cctcgcggg ccacacgcgc ccgggaccgc      2040
ggactaagcc tgtcgcccat gcccgcgcgc ccccgagca acgcgctggc caacgacggc      2100
cgcacgaacg tcgcgcctc cagcgccctg ttgacgaag tcaaacgcgg ccgacaccag      2160
agccattaa

```

<210> 225

<211> 1765

<212> DNA

<213> Herpes simplex virus

<220>

<221> misc_feature

<222> (1)...(1765)

<223> n = A,T,C or G

<400> 225

```

cccgatggac ccacgcgacg ttttggggcg ggtgggcggt tcgcggtggt tgccctcgcc      60
gctgttctcg gacgagctca gctacgagga ggacgactac cccgcgcgcg tcgcgcacga      120
tgacgncgcc gggcgcgccg ctcgcgcgac ggtcgagatt ctgcggggcc gcgtgtcggg      180
cccgagctg caggcgccat tccccctgga ccgcctgacc ccccgagtcg ccgctgtgga      240

```


cgagtcggtg	cgctcggccc	tggccctggg	acatccggcc	gggttctacc	cggtgcggga	300
tagcgcgttc	gggctgtcgc	gcgtgggggt	catgcacttt	gcctccccgg	ccgacccaaa	360
ggtgtttttc	cgccagaacg	tgcagcaggg	cgaggcgctg	gcctggtagc	tcacggggcg	420
cgcgatcctc	gacntgaacg	atcggcgggc	aaaaaccagc	ccctccccgg	cgatgggctt	480
tctgggtggac	gccatcgtgc	gggtggcgat	caacgggtgg	gtctnnngga	cgcgctgca	540
cacggagggg	cgcggtcggg	agctcgacga	cagggcggcc	gagctccgac	ggcagttcgc	600
ganctcnnn	ncgnnnnngn	ncgtgggggc	cgccgcccgt	ccgctgctca	gcgcgggagg	660
ggcgcgcgcc	ccccaccccg	gccccgacgc	cgcggtcttt	cgagttcgc	tggggtccct	720
gctgtactgg	cccggggtgc	gcgcgctcct	ggggcgcgac	tgtcgcgtgg	ccgcccgcga	780
cgcgggggcg	atgacgtaca	tgcacacccg	ggctctgctc	gcccgttca	accccgggcg	840
cgtcaaatgc	gtgtccccgc	gggaggccgc	gtttgcgggn	cgctcctgg	acgtgctggc	900
ggtcctggcg	gagnagacgg	tccagtggct	ctcggtggtc	gtggggcgcg	gcctgcaccc	960
gcactccgcc	caccccgctg	ttgcggacgt	ggagcaggag	gcgtgtttc	gcgccctgcc	1020
cctgggtagc	cccggggtcg	tggcgggcga	gcacgaggcg	ctgggcgaca	ccgcccgcgc	1080
ccgctgtctc	gccaccagcg	ggctgaacgc	cgtgctgggc	gcggccgtgt	acgcgctgca	1140
cacggccctg	gcgaccgtta	ccctgaaata	cgccctggcn	tgcggggacg	cnngccggcg	1200
caggagcagc	gcggcgcccg	cgcgcgccgt	gctggcgacg	gggctcatcc	tgcagcggct	1260
gctgggcctg	gcccacacgg	tggtcgcgtg	cgtggccctg	gcccgtttg	acggcgggtc	1320
gacggccccc	gaggtgggca	cgtaaccccc	cctgcgctac	gcgtgcgtcc	tccgcgcgac	1380
ccagcccttg	tacgcgcgga	ccacccccgc	caaatttttg	gcggacgtgc	gcgcgcgcgc	1440
ggaacacgtg	gaccttcgcc	ccgcgtcctc	ggcgccccgg	gcgcccgtga	gcgggacggc	1500
agaccccgcc	ttcctgctcg	aagacctggc	ggccttcccc	cccgcggccc	tgaatagcga	1560
gtcgtgctg	gggcgcgggg	tccgcgtcgt	ggacatcatg	gcgcagtttc	ggaaactgct	1620
catgggcgac	gaggagaccg	ccgcctcccg	ggcgacgtg	tccgggaggc	gcgcgaccgg	1680
gctgggcggc	ccgccacgcc	cataggcgcc	tccccataaa	aagcaacct	atatcccggg	1740
acggggcata	cctccgaccg	gcggg				1765

<210> 226

<211> 2091

<212> DNA

<213> Herpes simplex virus

<400> 226

atgtccgtgc	gcgggcatgc	cgtacgcggg	aggcgcgctt	ccacccggtc	ccatgccccg	60
tccgcgcata	gcgcgcactc	gcccgtggag	gacgagcccg	agggcggtgg	agtcgggtta	120
atgggggtacc	tgcggggcgt	gtttaacgtg	gacgacgaca	gcgaggtcga	ggccgcgggg	180
gagatggcga	gcgaagagcc	gccccgcgc	cgctcgccggg	agggcccgcg	tcaccccggg	240
tcccagacgc	cgtccgaggg	ccgggcggcg	gcgcggccccc	gcccggcgctc	ctttccgcgc	300
cccagggtccg	ttacggccag	gagccagtcc	gttcgcggac	gcccgggacag	cgccatcacg	360
cgggccccgc	ggggaggcta	cctgggcccg	atggacccac	gcgacgtttt	ggggcggtg	420
ggcggttcgc	gggtgggtgc	ctcgccgctg	ttcctggacg	agctcaacta	cgaggaggac	480
gactaccccc	ccgcgcgtgc	gcacgatgac	ggccccgggg	cgcgcccttc	cgcgacggtc	540
gagattctcg	cgggccgcgt	gtcgggcccg	gagctgcagg	cggcattccc	cctggaccgc	600
ctgaccccc	gagtcgcgcg	gtgggacgag	tccgtgcgct	cggccctggc	cctgggacat	660
ccggccgggt	tctacccgtg	tccggatagc	gcgttcgggc	tgtcgcgcgt	gggggtcatg	720
cactttgcct	ccccggccga	cccaaagggtg	tttttcgcgc	agacgctgca	gcagggcgag	780
gcgctggcct	ggtacgtcac	gggcgacgcg	atcctcgacc	tgacggatcg	gcgggcaaaa	840
accagccct	ccgcgcgat	gggttttctg	gtggacgcca	tcgtgcgggt	ggcgatcaac	900
gggtgggtct	gcgggacgcg	cctgcacacg	gaggggcgcg	gctcggagct	cgacgacagg	960
gcggccgagc	tccgacggca	gttcgcgagc	ctcacggcgt	tgcggcccg	gggggcgcgc	1020
gcggtgcgc	tgtcagcgc	gggagggggc	gcgcggccccc	accccgggcc	cgacgcgcgc	1080
gtcttttcga	gttcgctggg	gtccctgctg	tactggcccg	gggtgcgcgc	gctcctgggg	1140
cgcgactgtc	gcgtggccgc	ccgctacgcg	gggcgcgatga	cgtacatcgc	caccggggct	1200
ctgctcgccc	gcttcaaccc	cggcgcgcgc	aaatgcgtgc	tcccgcggga	ggccgcgttt	1260
gcggggcgcg	tcctggacgt	gctggcggtc	ctggcgagac	agacggtcca	gtggtctctg	1320
gtggtcgtgg	gggcgcgcct	gcacccgcac	tccgcccacc	ccgcgtttgc	ggacgtggag	1380
caggagcgcg	tgtttcgcgc	cctgcccctg	ggcagcccg	gggtcgtggc	ggccgagcac	1440
gaggcgctgg	gcgacaccgc	ggcgcgccgc	ctgctcgcca	ccagcgggct	gaacgcgcgtg	1500

```

ctgggcgcgcg cegtgtacgc gctgcacacg gccctggcgga cegtaccct gaaatacgcc 1560
ctggcctgcg gggacgcgcg ccggcgcgag gacgacgcgg cgccgcgcg cgccgtgctg 1620
gcgacggggc tcacctcgca gcggctgctg gccctggccg acacggtggt cgcgtgcgtg 1680
gccctggccg cgtttgacgg cggttcgacg gcccccgagg tgggcacgta cccccctg 1740
cgctacgcgt gcgtcctccg cgcgacccag cccctgtacg cgcgaccac ccccgccaaa 1800
ttttgggcgg acgtgcgcgc cgccgcggaa cacgtggacc ttgcggccgc gtcctcggcg 1860
ccccggggcg ccgtgagcgg gacggcagac cccgccttcc tgctcgaaga cctggcgcc 1920
ttcccccccg cccccctgaa tagcgagtcg gtgctggggc cgcggtccg cgtcgtggac 1980
atcatggcgc agtttcggaa actgctcatg ggcgacgagg agaccgccgc cctccgggcg 2040
cacgtgtccg ggaggcgcg gaccgggctg ggcggcccgc cacgcccata g 2091

```

<210> 227

<211> 292

<212> PRT

<213> Herpes simplex virus

<400> 227

```

Val Gln Val Lys His Ile Asp Arg Val Val Ser Pro Ser Val Ser Ser
 1          5          10          15
Ala Pro Pro Pro Ser Ala Pro Asp Ala Ser Leu Pro Pro Pro Gly Leu
 20          25          30
Gln Glu Ala Ala Pro Pro Gly Pro Pro Leu Arg Glu Leu Trp Trp Val
 35          40          45
Phe Tyr Ala Gly Asp Arg Ala Leu Glu Glu Pro His Ala Glu Ser Gly
 50          55          60
Leu Thr Arg Glu Glu Val Arg Ala Val His Gly Phe Arg Glu Gln Ala
 65          70          75          80
Trp Lys Leu Phe Gly Ser Val Gly Ala Pro Arg Ala Phe Leu Gly Ala
 85          90          95
Ala Leu Ala Leu Ser Pro Thr Gln Lys Leu Ala Val Tyr Tyr Tyr Leu
100          105          110
Ile His Arg Glu Arg Arg Met Ser Pro Phe Pro Ala Leu Val Arg Leu
115          120          125
Val Gly Arg Tyr Ile Gln Arg His Gly Leu Tyr Val Pro Ala Pro Asp
130          135          140
Glu Pro Thr Leu Ala Asp Ala Met Asn Gly Leu Phe Arg Asp Ala Leu
145          150          155          160
Ala Ala Gly Thr Val Ala Glu Gln Leu Leu Met Phe Asp Leu Leu Pro
165          170          175
Pro Lys Asp Val Pro Val Gly Ser Asp Ala Arg Ala Asp Ser Ala Ala
180          185          190
Leu Leu Arg Phe Val Asp Ser Gln Arg Leu Thr Pro Gly Gly Ser Val
195          200          205
Ser Pro Glu His Val Met Tyr Leu Gly Ala Phe Leu Gly Val Leu Tyr
210          215          220
Ala Gly His Gly Arg Leu Ala Ala Ala Thr His Thr Ala Arg Leu Thr
225          230          235          240
Gly Val Thr Ser Leu Val Leu Thr Val Gly Asp Val Asp Arg Met Ser
245          250          255
Ala Phe Asp Arg Gly Pro Ala Gly Ala Ala Gly Arg Thr Arg Thr Ala
260          265          270
Gly Tyr Leu Asp Ala Leu Leu Thr Val Cys Leu Ala Arg Ala Gln His
275          280          285
Gly Gln Ser Val
290

```

<210> 228

<211> 532

<212> PRT

<213> Herpes simplex virus

<400> 228

```

Met Glu Leu Ser Tyr Ala Thr Thr Leu His His Arg Asp Val Val Phe
 1          5          10          15
Tyr Val Thr Ala Asp Arg Asn Arg Ala Tyr Phe Val Cys Gly Gly Ser
      20          25          30
Val Tyr Ser Val Gly Arg Pro Arg Asp Ser Gln Pro Gly Glu Ile Ala
      35          40          45
Lys Phe Gly Leu Val Val Arg Gly Thr Gly Pro Lys Asp Arg Met Val
      50          55          60
Ala Asn Tyr Val Arg Ser Glu Leu Arg Gln Arg Gly Leu Arg Asp Val
65          70          75          80
Arg Pro Val Gly Glu Asp Glu Val Phe Leu Asp Ser Val Cys Leu Leu
      85          90          95
Asn Pro Asn Val Ser Ser Glu Arg Asp Val Ile Asn Thr Asn Asp Val
      100          105          110
Glu Val Leu Asp Glu Cys Leu Ala Glu Tyr Cys Thr Ser Leu Arg Thr
      115          120          125
Ser Pro Gly Val Leu Val Thr Gly Val Arg Val Arg Ala Arg Asp Arg
      130          135          140
Val Ile Glu Leu Phe Glu His Pro Ala Ile Val Asn Ile Ser Ser Arg
      145          150          155          160
Phe Ala Tyr Thr Pro Ser Pro Tyr Val Phe Ala Leu Ala Gln Ala His
      165          170          175
Leu Pro Arg Leu Pro Ser Ser Leu Glu Pro Leu Val Ser Gly Leu Phe
      180          185          190
Asp Gly Ile Pro Ala Pro Arg Gln Pro Leu Asp Ala Arg Asp Arg Arg
      195          200          205
Thr Asp Val Val Ile Thr Gly Thr Arg Ala Pro Arg Pro Met Ala Gly
      210          215          220
Thr Gly Ala Gly Gly Ala Gly Ala Lys Arg Ala Thr Val Ser Glu Phe
      225          230          235          240
Val Gln Val Lys His Ile Asp Arg Val Val Ser Pro Ser Val Ser Ser
      245          250          255
Ala Pro Pro Pro Ser Ala Pro Asp Ala Ser Leu Pro Pro Pro Gly Leu
      260          265          270
Gln Glu Ala Ala Pro Pro Gly Pro Pro Leu Arg Glu Leu Trp Trp Val
      275          280          285
Phe Tyr Ala Gly Asp Arg Ala Leu Glu Glu Pro His Ala Glu Ser Gly
      290          295          300
Leu Thr Arg Glu Glu Val Arg Ala Val His Gly Phe Arg Glu Gln Ala
      305          310          315          320
Trp Lys Leu Phe Gly Ser Val Gly Ala Pro Arg Ala Phe Leu Gly Ala
      325          330          335
Ala Leu Ala Leu Ser Pro Thr Gln Lys Leu Ala Val Tyr Tyr Tyr Leu
      340          345          350
Ile His Arg Glu Arg Arg Met Ser Pro Phe Pro Ala Leu Val Arg Leu
      355          360          365
Val Gly Arg Tyr Ile Gln Arg His Gly Leu Tyr Val Pro Ala Pro Asp
      370          375          380
Glu Pro Thr Leu Ala Asp Ala Met Asn Gly Leu Phe Arg Asp Ala Leu
      385          390          395          400
Ala Ala Gly Thr Val Ala Glu Gln Leu Leu Met Phe Asp Leu Leu Pro
      405          410          415
Pro Lys Asp Val Pro Val Gly Ser Asp Ala Arg Ala Asp Ser Ala Ala
      420          425          430

```

Leu Leu Arg Phe Val Asp Ser Gln Arg Leu Thr Pro Gly Gly Ser Val
 435 440 445
 Ser Pro Glu His Val Met Tyr Leu Gly Ala Phe Leu Gly Val Leu Tyr
 450 455 460
 Ala Gly His Gly Arg Leu Ala Ala Ala Thr His Thr Ala Arg Leu Thr
 465 470 475 480
 Gly Val Thr Ser Leu Val Leu Thr Val Gly Asp Val Asp Arg Met Ser
 485 490 495
 Ala Phe Asp Arg Gly Pro Ala Gly Ala Ala Gly Arg Thr Arg Thr Ala
 500 505 510
 Gly Tyr Leu Asp Ala Leu Leu Thr Val Cys Leu Ala Arg Ala Gln His
 515 520 525
 Gly Gln Ser Val
 530

<210> 229

<211> 20

<212> PRT

<213> Herpes simplex virus

<400> 229

Pro Leu Arg Glu Leu Trp Trp Val Phe Tyr Ala Gly Asp Arg Ala Leu
 1 5 10 15
 Glu Glu Pro His
 20

<210> 230

<211> 248

<212> PRT

<213> Herpes simplex virus

<400> 230

Val Ala Pro Leu Ile Val Gly Leu Thr Val Glu Ser Gly Phe Glu Ala
 1 5 10 15
 Asn Val Ala Ala Val Val Gly Ser Arg Thr Thr Gly Leu Gly Gly Thr
 20 25 30
 Ala Val Ser Leu Lys Leu Met Pro Ser His Tyr Ser Pro Ser Val Tyr
 35 40 45
 Val Phe His Gly Gly Arg His Leu Ala Pro Ser Thr Gln Ala Pro Asn
 50 55 60
 Leu Thr Arg Leu Cys Glu Arg Ala Arg Arg His Phe Gly Phe Ser Asp
 65 70 75 80
 Tyr Ala Pro Arg Pro Cys Asp Leu Lys His Glu Thr Thr Gly Asp Ala
 85 90 95
 Leu Cys Glu Arg Leu Gly Leu Asp Pro Asp Arg Ala Leu Leu Tyr Leu
 100 105 110
 Val Ile Thr Glu Gly Phe Arg Glu Ala Val Cys Ile Ser Asn Thr Phe
 115 120 125
 Leu His Leu Gly Gly Met Asp Lys Val Thr Ile Gly Asp Ala Glu Val
 130 135 140
 His Arg Ile Pro Val Tyr Pro Leu Gln Met Phe Met Pro Asp Phe Ser
 145 150 155 160
 Arg Val Ile Ala Asp Pro Phe Asn Cys Asn His Arg Ser Ile Gly Glu
 165 170 175
 Asn Phe Asn Tyr Pro Leu Pro Phe Phe Asn Arg Pro Leu Ala Arg Leu
 180 185 190
 Leu Phe Glu Ala Val Val Gly Pro Ala Ala Val Ala Leu Arg Ala Arg
 195 200 205

Asn Val Asp Ala Val Ala Arg Ala Ala Ala His Leu Ala Phe Asp Glu
 210 215 220
 Asn His Glu Gly Ala Ala Leu Pro Ala Asp Ile Thr Phe Thr Ala Phe
 225 230 235 240
 Glu Ala Ser Gln Gly Lys Pro Gln
 245

<210> 231
 <211> 274
 <212> PRT
 <213> Herpes simplex virus

<400> 231
 Gly Leu Glu Leu Leu Ser Leu Leu Ser Ala Arg Ser Gly Asp Ala Asp
 1 5 10 15
 Val Ala Val Ala Pro Leu Ile Val Gly Leu Thr Val Glu Ser Gly Phe
 20 25 30
 Glu Ala Asn Val Ala Ala Val Val Gly Ser Arg Thr Thr Gly Leu Gly
 35 40 45
 Gly Thr Ala Val Ser Leu Lys Leu Met Pro Ser His Tyr Ser Pro Ser
 50 55 60
 Val Tyr Val Phe His Gly Gly Arg His Leu Ala Pro Ser Thr Gln Ala
 65 70 75 80
 Pro Asn Leu Thr Arg Leu Cys Glu Arg Ala Arg Arg His Phe Gly Phe
 85 90 95
 Ser Asp Tyr Ala Pro Arg Pro Cys Asp Leu Lys His Glu Thr Thr Gly
 100 105 110
 Asp Ala Leu Cys Glu Arg Leu Gly Leu Asp Pro Asp Arg Ala Leu Leu
 115 120 125
 Tyr Leu Val Ile Thr Glu Gly Phe Arg Glu Ala Val Cys Ile Ser Asn
 130 135 140
 Thr Phe Leu His Leu Gly Gly Met Asp Lys Val Thr Ile Gly Asp Ala
 145 150 155 160
 Glu Val His Arg Ile Pro Val Tyr Pro Leu Gln Met Phe Met Pro Asp
 165 170 175
 Phe Ser Arg Val Ile Ala Asp Pro Phe Asn Cys Asn His Arg Ser Ile
 180 185 190
 Gly Glu Asn Phe Asn Tyr Pro Leu Pro Phe Phe Asn Arg Pro Leu Ala
 195 200 205
 Arg Leu Leu Phe Glu Ala Val Val Gly Pro Ala Ala Val Ala Leu Arg
 210 215 220
 Ala Arg Asn Val Asp Ala Val Ala Arg Ala Ala His Leu Ala Phe
 225 230 235 240
 Asp Glu Asn His Glu Gly Ala Ala Leu Pro Ala Asp Ile Thr Phe Thr
 245 250 255
 Ala Phe Glu Ala Ser Gln Gly Lys Pro Gln Arg Gly Ala Arg Asp Ala
 260 265 270
 Gly Asn

<210> 232
 <211> 1196
 <212> PRT
 <213> Herpes simplex virus

<400> 232
 Met Asp Thr Lys Pro Lys Thr Thr Thr Thr Val Lys Val Pro Pro Gly
 1 5 10 15

Pro Met Gly Tyr Val Tyr Gly Arg Ala Cys Pro Ala Glu Gly Leu Glu
 20 25 30
 Leu Leu Ser Leu Leu Ser Ala Arg Ser Gly Asp Ala Asp Val Ala Val
 35 40 45
 Ala Pro Leu Ile Val Gly Leu Thr Val Glu Ser Gly Phe Glu Ala Asn
 50 55 60
 Val Ala Ala Val Val Gly Ser Arg Thr Thr Gly Leu Gly Gly Thr Ala
 65 70 75 80
 Val Ser Leu Lys Leu Met Pro Ser His Tyr Ser Pro Ser Val Tyr Val
 85 90 95
 Phe His Gly Gly Arg His Leu Ala Pro Ser Thr Gln Ala Pro Asn Leu
 100 105 110
 Thr Arg Leu Cys Glu Arg Ala Arg Pro His Phe Gly Phe Ala Asp Tyr
 115 120 125
 Ala Pro Arg Pro Cys Asp Leu Lys His Glu Thr Thr Gly Asp Ala Leu
 130 135 140
 Cys Glu Arg Leu Gly Leu Asp Pro Asp Arg Ala Leu Leu Tyr Leu Val
 145 150 155 160
 Ile Thr Glu Gly Phe Arg Glu Ala Val Cys Ile Ser Asn Thr Phe Leu
 165 170 175
 His Leu Gly Gly Met Asp Lys Val Thr Ile Gly Asp Ala Glu Val His
 180 185 190
 Arg Ile Pro Val Tyr Pro Leu Gln Met Phe Met Pro Asp Phe Ser Arg
 195 200 205
 Val Ile Ala Asp Pro Phe Asn Cys Asn His Arg Ser Ile Gly Glu Asn
 210 215 220
 Phe Asn Tyr Pro Leu Pro Phe Phe Asn Arg Pro Leu Ala Arg Leu Leu
 225 230 235 240
 Phe Glu Ala Val Val Gly Pro Ala Ala Val Ala Leu Arg Ala Arg Asn
 245 250 255
 Val Asp Ala Val Ala Arg Ala Ala Ala His Leu Ala Phe Asp Glu Asn
 260 265 270
 His Glu Gly Ala Ala Leu Pro Ala Asp Ile Thr Phe Thr Ala Phe Glu
 275 280 285
 Ala Ser Gln Gly Lys Pro Gln Arg Gly Ala Arg Asp Ala Gly Asn Lys
 290 295 300
 Gly Pro Ala Gly Gly Phe Glu Gln Arg Leu Ala Ser Val Met Ala Gly
 305 310 315 320
 Asp Ala Ala Leu Ala Leu Glu Ser Ile Val Ser Met Ala Val Phe Asp
 325 330 335
 Glu Pro Pro Pro Asp Ile Thr Thr Trp Pro Leu Leu Glu Gly Gln Glu
 340 345 350
 Thr Pro Ala Ala Arg Ala Gly Ala Val Gly Ala Tyr Leu Ala Arg Ala
 355 360 365
 Ala Gly Leu Val Gly Ala Met Val Phe Ser Thr Asn Ser Ala Leu His
 370 375 380
 Leu Thr Glu Val Asp Asp Ala Gly Pro Ala Asp Pro Lys Asp His Ser
 385 390 395 400
 Lys Pro Ser Phe Tyr Arg Phe Phe Leu Val Pro Gly Thr His Val Ala
 405 410 415
 Ala Asn Pro Gln Leu Asp Arg Glu Gly His Val Val Pro Gly Tyr Glu
 420 425 430
 Gly Arg Pro Thr Ala Pro Leu Val Gly Gly Thr Gln Glu Phe Ala Gly
 435 440 445
 Glu His Leu Ala Met Leu Cys Gly Phe Ser Pro Ala Leu Leu Ala Lys
 450 455 460
 Met Leu Phe Tyr Leu Glu Arg Cys Asp Gly Gly Val Ile Val Gly Arg
 465 470 475 480

Gln Glu Met Asp Val Phe Arg Tyr Val Ala Asp Ser Gly Gln Thr Asp
 485 490 495
 Val Pro Cys Asn Leu Cys Thr Phe Glu Thr Arg His Ala Cys Ala His
 500 505 510
 Thr Thr Leu Met Arg Leu Arg Ala Arg His Pro Lys Phe Ala Ser Ala
 515 520 525
 Ala Arg Gly Ala Ile Gly Val Phe Gly Thr Met Asn Ser Ala Tyr Ser
 530 535 540
 Asp Cys Asp Val Leu Gly Asn Tyr Ala Ala Phe Ser Ala Leu Lys Arg
 545 550 555 560
 Ala Asp Gly Ser Glu Asn Thr Arg Thr Ile Met Gln Glu Thr Tyr Arg
 565 570 575
 Ala Ala Thr Glu Arg Val Met Ala Glu Leu Glu Ala Leu Gln Tyr Val
 580 585 590
 Asp Gln Ala Val Pro Thr Ala Leu Gly Arg Leu Glu Thr Ile Ile Gly
 595 600 605
 Asn Arg Glu Ala Leu His Thr Val Val Asn Asn Ile Lys Gln Leu Val
 610 615 620
 Asp Arg Glu Val Glu Gln Leu Met Arg Asn Leu Ile Glu Gly Arg Asn
 625 630 635 640
 Phe Lys Phe Arg Asp Gly Leu Ala Glu Ala Asn His Ala Met Ser Leu
 645 650 655
 Ser Leu Asp Pro Tyr Thr Cys Gly Pro Cys Pro Leu Leu Gln Leu Leu
 660 665 670
 Ala Arg Arg Ser Asn Leu Ala Val Tyr Gln Asp Leu Ala Leu Ser Gln
 675 680 685
 Cys His Gly Val Phe Ala Gly Gln Ser Val Glu Gly Arg Asn Phe Arg
 690 695 700
 Asn Gln Phe Gln Pro Val Leu Arg Arg Arg Val Met Asp Leu Phe Asn
 705 710 715 720
 Asn Gly Phe Leu Ser Ala Lys Thr Leu Thr Val Ala Leu Ser Glu Gly
 725 730 735
 Ala Ala Ile Cys Ala Pro Ser Leu Thr Ala Gly Gln Thr Ala Pro Ala
 740 745 750
 Glu Ser Ser Phe Glu Gly Asp Val Ala Arg Val Thr Leu Gly Phe Pro
 755 760 765
 Lys Glu Leu Arg Val Lys Ser Arg Val Leu Phe Ala Gly Ala Ser Ala
 770 775 780
 Asn Ala Ser Glu Ala Ala Lys Ala Arg Val Ala Ser Leu Gln Ser Ala
 785 790 795 800
 Tyr Gln Lys Pro Asp Lys Arg Val Asp Ile Leu Leu Gly Pro Leu Gly
 805 810 815
 Phe Leu Leu Lys Gln Phe His Ala Val Ile Phe Pro Asn Gly Lys Pro
 820 825 830
 Pro Gly Ser Asn Gln Pro Asn Pro Gln Trp Phe Trp Thr Ala Leu Gln
 835 840 845
 Arg Asn Gln Leu Pro Ala Arg Leu Leu Ser Arg Glu Asp Ile Glu Thr
 850 855 860
 Ile Ala Phe Ile Lys Arg Phe Ser Leu Asp Tyr Gly Ala Ile Asn Phe
 865 870 875 880
 Ile Asn Leu Ala Pro Asn Asn Val Ser Glu Leu Ala Met Tyr Tyr Met
 885 890 895
 Ala Asn Gln Ile Leu Arg Tyr Cys Asp His Ser Thr Tyr Phe Ile Asn
 900 905 910
 Thr Leu Thr Ala Val Ile Ala Gly Ser Arg Arg Pro Pro Ser Val Gln
 915 920 925
 Ala Ala Ala Ala Trp Ala Pro Gln Gly Gly Ala Gly Leu Glu Ala Gly
 930 935 940

Ala Arg Ala Leu Met Asp Ser Leu Asp Ala His Pro Gly Ala Trp Thr
 945 950 955 960
 Ser Met Phe Ala Ser Cys Asn Leu Leu Arg Pro Val Met Ala Ala Arg
 965 970 975
 Pro Met Val Val Leu Gly Leu Ser Ile Ser Lys Tyr Tyr Gly Met Ala
 980 985 990
 Gly Asn Asp Arg Val Phe Gln Ala Gly Asn Trp Ala Ser Leu Leu Gly
 995 1000 1005
 Gly Lys Asn Ala Cys Pro Leu Leu Ile Phe Asp Arg Thr Arg Lys Phe
 1010 1015 1020
 Val Leu Ala Cys Pro Arg Ala Gly Phe Val Cys Ala Ala Ser Ser Leu
 1025 1030 1035 1040
 Gly Gly Gly Ala His Glu His Ser Leu Cys Glu Gln Leu Arg Gly Ile
 1045 1050 1055
 Ile Ala Glu Gly Gly Ala Ala Val Ala Ser Ser Val Phe Val Ala Thr
 1060 1065 1070
 Val Lys Ser Leu Gly Pro Arg Thr Gln Gln Leu Gln Ile Glu Asp Trp
 1075 1080 1085
 Leu Ala Leu Leu Glu Asp Glu Tyr Leu Ser Glu Glu Met Met Glu Phe
 1090 1095 1100
 Thr Thr Arg Ala Leu Glu Arg Gly His Gly Glu Trp Ser Thr Asp Ala
 1105 1110 1115 1120
 Ala Leu Glu Val Ala His Glu Ala Glu Ala Leu Val Ser Gln Leu Gly
 1125 1130 1135
 Ala Ala Gly Glu Val Phe Asn Phe Gly Asp Phe Gly Asp Glu Asp Asp
 1140 1145 1150
 His Ala Ala Ser Phe Gly Gly Leu Ala Ala Ala Ala Gly Ala Ala Gly
 1155 1160 1165
 Val Ala Arg Lys Arg Ala Phe His Gly Asp Asp Pro Phe Gly Glu Gly
 1170 1175 1180
 Pro Pro Glu Lys Lys Asp Leu Thr Leu Asp Met Leu
 1185 1190 1195

<210> 233

<211> 193

<212> PRT

<213> Herpes simplex virus

<400> 233

Arg Pro Leu Ala Ala Gln Arg Glu Ala Ala Gly Val Tyr Asp Ala Val
 1 5 10 15
 Arg Thr Trp Gly Pro Asp Ala Glu Ala Glu Pro Asp Gln Met Glu Asn
 20 25 30
 Thr Tyr Leu Leu Pro Asp Asp Asp Ala Ala Met Pro Ala Gly Val Gly
 35 40 45
 Leu Gly Ala Thr Pro Ala Ala Asp Thr Thr Ala Ala Trp Pro Ala
 50 55 60
 Glu Ser His Ala Pro Arg Ala Pro Ser Glu Asp Ala Asp Ser Ile Tyr
 65 70 75 80
 Glu Ser Val Ser Glu Asp Gly Gly Arg Val Tyr Glu Glu Ile Pro Trp
 85 90 95
 Val Arg Val Tyr Glu Asn Ile Cys Leu Arg Arg Gln Asp Ala Gly Gly
 100 105 110
 Ala Ala Pro Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala Glu
 115 120 125
 Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro Pro Gly
 130 135 140
 Ala Thr Arg Ala Pro Asp Pro Gly Leu Ser Leu Ser Pro Met Pro Ala

145 150 155 160
 Arg Pro Arg Thr Asn Ala Leu Ala Asn Asp Gly Pro Thr Asn Val Ala
 165 170 175
 Ala Leu Ser Ala Leu Leu Thr Lys Leu Lys Arg Gly Arg His Gln Ser
 180 185 190
 His

<210> 234
 <211> 277
 <212> PRT
 <213> Herpes simplex virus

<400> 234
 Arg Thr Arg Arg Asp Asn Ala Pro Pro Pro Leu Pro Arg Ala Arg Pro
 1 5 10 15
 His Ser Thr Pro Ala Ala Pro Arg Arg Cys Arg Arg His Arg Glu Asp
 20 25 30
 Leu Pro Glu Pro Pro His Val Asp Ala Ala Asp Arg Gly Pro Glu Pro
 35 40 45
 Cys Ala Gly Arg Pro Ala Thr Tyr Tyr Thr His Met Ala Gly Ala Pro
 50 55 60
 Pro Arg Leu Pro Pro Arg Asn Pro Ala Pro Pro Glu Gln Arg Pro Ala
 65 70 75 80
 Ala Ala Ala Arg Pro Leu Ala Ala Gln Arg Glu Ala Ala Gly Val Tyr
 85 90 95
 Asp Ala Val Arg Thr Trp Gly Pro Asp Ala Glu Ala Glu Pro Asp Gln
 100 105 110
 Met Glu Asn Thr Tyr Leu Leu Pro Asp Asp Asp Ala Ala Met Pro Ala
 115 120 125
 Gly Val Gly Leu Gly Ala Thr Pro Ala Ala Asp Thr Thr Ala Ala Ala
 130 135 140
 Ala Trp Pro Ala Glu Ser His Ala Pro Arg Ala Pro Ser Glu Asp Ala
 145 150 155 160
 Asp Ser Ile Tyr Glu Ser Val Gly Glu Asp Gly Gly Arg Val Tyr Glu
 165 170 175
 Glu Ile Pro Trp Val Arg Val Tyr Glu Asn Ile Cys Pro Arg Arg Arg
 180 185 190
 Leu Ala Gly Gly Ala Ala Leu Pro Gly Asp Ala Pro Asp Ser Pro Tyr
 195 200 205
 Ile Glu Ala Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe
 210 215 220
 Ser Pro Arg Arg Ala Thr Arg Ala Pro Asp Pro Gly Leu Ser Leu Ser
 225 230 235 240
 Pro Met Pro Ala Arg Pro Arg Thr Asn Ala Leu Ala Asn Asp Gly Pro
 245 250 255
 Thr Asn Val Ala Ala Leu Ser Ala Leu Leu Thr Lys Leu Lys Arg Gly
 260 265 270
 Arg His Gln Ser His
 275

<210> 235
 <211> 722
 <212> PRT
 <213> Herpes simplex virus

<400> 235
 Met Gln Arg Arg Ala Arg Gly Ala Ser Ser Leu Arg Leu Ala Arg Cys

1	5	10	15
Leu Thr Pro Ala	Asn Leu Ile Arg Gly Ala Asn Ala Gly Val Pro Glu		
20		25	30
Arg Arg Ile Phe Ala Gly Cys Leu Pro Thr Pro Glu Gly Leu Leu			
35	40	45	
Ser Ala Ala Val Gly Val Leu Arg Gln Arg Ala Asp Asp Leu Gln Pro			
50	55	60	
Ala Phe Leu Thr Gly Ala Asp Arg Ser Val Arg Leu Ala Ala Arg His			
65	70	75	80
His Asn Thr Val Pro Glu Ser Leu Ile Val Asp Gly Leu Ala Ser Asp			
85	90	95	
Pro His Tyr Asp Tyr Ile Arg His Tyr Ala Ser Ala Ala Lys Gln Ala			
100	105	110	
Leu Gly Glu Val Glu Leu Ser Gly Gln Leu Ser Arg Ala Ile Leu			
115	120	125	
Ala Gln Tyr Trp Lys Tyr Leu Gln Thr Val Val Pro Ser Gly Leu Asp			
130	135	140	
Ile Pro Asp Asp Pro Ala Gly Asp Cys Asp Pro Ser Leu His Val Leu			
145	150	155	160
Leu Arg Pro Thr Leu Leu Pro Lys Leu Leu Val Arg Ala Pro Phe Lys			
165	170	175	
Ser Gly Ala Ala Ala Ala Lys Tyr Ala Ala Ala Val Ala Gly Leu Arg			
180	185	190	
Asp Ala Ala His Arg Leu Gln Gln Tyr Met Phe Phe Met Arg Pro Ala			
195	200	205	
Asp Pro Ser Arg Pro Ser Thr Asp Thr Ala Leu Arg Leu Ser Glu Leu			
210	215	220	
Leu Ala Tyr Val Ser Val Leu Tyr His Trp Ala Ser Trp Met Leu Trp			
225	230	235	240
Thr Ala Asp Lys Tyr Val Cys Arg Arg Leu Gly Pro Ala Asp Arg Arg			
245	250	255	
Phe Val Ala Leu Ser Gly Ser Leu Glu Ala Pro Ala Glu Thr Phe Ala			
260	265	270	
Arg His Leu Asp Arg Gly Pro Ser Gly Thr Thr Gly Ser Met Gln Cys			
275	280	285	
Met Ala Leu Arg Ala Ala Val Ser Asp Val Leu Gly His Leu Thr Arg			
290	295	300	
Leu Ala His Leu Trp Glu Thr Gly Lys Arg Ser Gly Gly Thr Tyr Gly			
305	310	315	320
Ile Val Asp Ala Ile Val Ser Thr Val Glu Val Leu Ser Ile Val His			
325	330	335	
His His Ala Gln Tyr Ile Ile Asn Ala Thr Leu Thr Gly Tyr Val Val			
340	345	350	
Trp Ala Ser Asp Ser Leu Asn Asn Glu Tyr Leu Thr Ala Ala Val Asp			
355	360	365	
Ser Gln Glu Arg Phe Cys Arg Thr Ala Ala Pro Leu Phe Pro Thr Met			
370	375	380	
Thr Ala Pro Ser Trp Ala Arg Met Glu Leu Ser Ile Lys Ser Trp Phe			
385	390	395	400
Gly Ala Ala Leu Ala Pro Asp Leu Leu Arg Ser Gly Thr Pro Ser Pro			
405	410	415	
His Tyr Glu Ser Ile Leu Arg Leu Ala Ala Ser Gly Pro Pro Gly Gly			
420	425	430	
Arg Gly Ala Val Gly Gly Ser Cys Arg Asp Lys Ile Gln Arg Thr Arg			
435	440	445	
Arg Asp Asn Ala Pro Pro Pro Leu Pro Arg Ala Arg Pro His Ser Thr			
450	455	460	
Pro Ala Ala Pro Arg Arg Cys Arg Arg His Arg Glu Asp Leu Pro Glu			

200

465 470 475 480
 Pro Pro His Val Asp Ala Ala Asp Arg Gly Pro Glu Pro Cys Ala Gly
 485 490 495
 Arg Pro Ala Thr Tyr Tyr Thr His Met Ala Gly Ala Pro Pro Arg Leu
 500 505 510
 Pro Pro Arg Asn Pro Ala Pro Pro Glu Gln Arg Pro Ala Ala Ala Ala
 515 520 525
 Arg Pro Leu Ala Ala Gln Arg Glu Ala Ala Gly Val Tyr Asp Ala Val
 530 535 540
 Arg Thr Trp Gly Pro Asp Ala Glu Ala Glu Pro Asp Gln Met Glu Asn
 545 550 555 560
 Thr Tyr Leu Leu Pro Asp Asp Asp Ala Ala Met Pro Ala Gly Val Gly
 565 570 575
 Leu Gly Ala Thr Pro Ala Ala Asp Thr Thr Ala Ala Ala Ala Trp Pro
 580 585 590
 Ala Glu Ser His Ala Pro Arg Ala Pro Ser Glu Asp Ala Asp Ser Ile
 595 600 605
 Tyr Glu Ser Val Gly Glu Asp Gly Gly Arg Val Tyr Glu Glu Ile Pro
 610 615 620
 Trp Val Arg Val Tyr Glu Asn Ile Cys Pro Arg Arg Arg Leu Ala Gly
 625 630 635 640
 Gly Ala Ala Leu Pro Gly Asp Ala Pro Asp Ser Pro Tyr Ile Glu Ala
 645 650 655
 Glu Asn Pro Leu Tyr Asp Trp Gly Gly Ser Ala Leu Phe Ser Pro Arg
 660 665 670
 Arg Ala Thr Arg Ala Pro Asp Pro Gly Leu Ser Leu Ser Pro Met Pro
 675 680 685
 Ala Arg Pro Arg Thr Asn Ala Leu Ala Asn Asp Gly Pro Thr Asn Val
 690 695 700
 Ala Ala Leu Ser Ala Leu Leu Thr Lys Leu Lys Arg Gly Arg His Gln
 705 710 715 720
 Ser His

<210> 236

<211> 19

<212> PRT

<213> Herpes simplex virus

<400> 236

Glu Glu Ile Pro Trp Val Arg Val Tyr Glu Asn Ile Cys Leu Arg Arg
 1 5 10 15
 Gln Asp Ala

<210> 237

<211> 567

<212> PRT

<213> Herpes simplex virus

<220>

<221> VARIANT

<222> (1)... (567)

<223> Xaa = Any Amino Acid

<400> 237

Pro Met Asp Pro Arg Asp Val Leu Gly Arg Val Gly Gly Ser Arg Val
 1 5 10 15

Val Pro Ser Pro Leu Phe Leu Asp Glu Leu Ser Tyr Glu Glu Asp Asp
 20 25 30
 Tyr Pro Ala Ala Val Ala His Asp Asp Xaa Ala Gly Ala Arg Pro Pro
 35 40 45
 Ala Thr Val Glu Ile Leu Ala Gly Arg Val Ser Gly Pro Glu Leu Gln
 50 55 60
 Ala Ala Phe Pro Leu Asp Arg Leu Thr Pro Arg Val Ala Ala Trp Asp
 65 70 75 80
 Glu Ser Val Arg Ser Ala Leu Ala Leu Gly His Pro Ala Gly Phe Tyr
 85 90 95
 Pro Cys Pro Asp Ser Ala Phe Gly Leu Ser Arg Val Gly Val Met His
 100 105 110
 Phe Ala Ser Pro Ala Asp Pro Lys Val Phe Phe Arg Gln Thr Leu Gln
 115 120 125
 Gln Gly Glu Ala Leu Ala Trp Tyr Val Thr Gly Asp Ala Ile Leu Asp
 130 135 140
 Xaa Thr Asp Arg Arg Ala Lys Thr Ser Pro Ser Arg Ala Met Gly Phe
 145 150 155 160
 Leu Val Asp Ala Ile Val Arg Val Ala Ile Asn Gly Trp Val Xaa Gly
 165 170 175
 Thr Arg Leu His Thr Glu Gly Arg Gly Ser Glu Leu Asp Asp Arg Ala
 180 185 190
 Ala Glu Leu Arg Arg Gln Phe Ala Xaa Leu Xaa Xaa Arg Xaa Val
 195 200 205
 Gly Ala Ala Val Pro Leu Ser Ala Gly Gly Ala Ala Pro Pro
 210 215 220
 His Pro Gly Pro Asp Ala Ala Val Phe Arg Ser Ser Leu Gly Ser Leu
 225 230 235 240
 Leu Tyr Trp Pro Gly Val Arg Ala Leu Leu Gly Arg Asp Cys Arg Val
 245 250 255
 Ala Ala Arg Tyr Ala Gly Arg Met Thr Tyr Ile Ala Thr Gly Ala Leu
 260 265 270
 Leu Ala Arg Phe Asn Pro Gly Ala Val Lys Cys Val Leu Pro Arg Glu
 275 280 285
 Ala Ala Phe Ala Gly Arg Val Leu Asp Val Leu Ala Val Leu Ala Glu
 290 295 300
 Xaa Thr Val Gln Trp Leu Ser Val Val Val Gly Ala Arg Leu His Pro
 305 310 315 320
 His Ser Ala His Pro Ala Phe Ala Asp Val Glu Gln Glu Ala Leu Phe
 325 330 335
 Arg Ala Leu Pro Leu Gly Ser Pro Gly Val Val Ala Ala Glu His Glu
 340 345 350
 Ala Leu Gly Asp Thr Ala Ala Arg Arg Leu Leu Ala Thr Ser Gly Leu
 355 360 365
 Asn Ala Val Leu Gly Ala Ala Val Tyr Ala Leu His Thr Ala Leu Ala
 370 375 380
 Thr Val Thr Leu Lys Tyr Ala Leu Ala Cys Gly Asp Ala Xaa Arg Arg
 385 390 395 400
 Arg Asp Asp Ala Ala Ala Ala Arg Ala Val Leu Ala Thr Gly Leu Ile
 405 410 415
 Leu Gln Arg Leu Leu Gly Leu Ala Asp Thr Val Val Ala Cys Val Ala
 420 425 430
 Leu Ala Ala Phe Asp Gly Gly Ser Thr Ala Pro Glu Val Gly Thr Tyr
 435 440 445
 Thr Pro Leu Arg Tyr Ala Cys Val Leu Arg Ala Thr Gln Pro Leu Tyr
 450 455 460
 Ala Arg Thr Thr Pro Ala Lys Phe Trp Ala Asp Val Arg Ala Ala Ala
 465 470 475 480

Glu His Val Asp Leu Arg Pro Ala Ser Ser Ala Pro Arg Ala Pro Val
 485 490 495
 Ser Gly Thr Ala Asp Pro Ala Phe Leu Leu Glu Asp Leu Ala Ala Phe
 500 505 510
 Pro Pro Ala Pro Leu Asn Ser Glu Ser Val Leu Gly Pro Arg Val Arg
 515 520 525
 Val Val Asp Ile Met Ala Gln Phe Arg Lys Leu Leu Met Gly Asp Glu
 530 535 540
 Glu Thr Ala Ala Xaa Arg Ala His Val Ser Gly Arg Arg Ala Thr Gly
 545 550 555 560
 Leu Gly Gly Pro Pro Arg Pro
 565

<210> 238

<211> 696

<212> PRT

<213> Herpes simplex virus

<400> 238

Met Ser Val Arg Gly His Ala Val Arg Arg Arg Arg Ala Ser Thr Arg
 1 5 10 15
 Ser His Ala Pro Ser Ala His Arg Ala Asp Ser Pro Val Glu Asp Glu
 20 25 30
 Pro Glu Gly Gly Gly Val Gly Leu Met Gly Tyr Leu Arg Ala Val Phe
 35 40 45
 Asn Val Asp Asp Asp Ser Glu Val Glu Ala Ala Gly Glu Met Ala Ser
 50 55 60
 Glu Glu Pro Pro Pro Arg Arg Arg Arg Glu Ala Arg Gly His Pro Gly
 65 70 75 80
 Ser Arg Arg Ala Ser Glu Ala Arg Ala Ala Ala Pro Pro Arg Arg Ala
 85 90 95
 Ser Phe Pro Arg Pro Arg Ser Val Thr Ala Arg Ser Gln Ser Val Arg
 100 105 110
 Gly Arg Arg Asp Ser Ala Ile Thr Arg Ala Pro Arg Gly Gly Tyr Leu
 115 120 125
 Gly Pro Met Asp Pro Arg Asp Val Leu Gly Arg Val Gly Gly Ser Arg
 130 135 140
 Val Val Pro Ser Pro Leu Phe Leu Asp Glu Leu Asn Tyr Glu Glu Asp
 145 150 155 160
 Asp Tyr Pro Ala Ala Val Ala His Asp Asp Gly Pro Gly Ala Arg Pro
 165 170 175
 Ser Ala Thr Val Glu Ile Leu Ala Gly Arg Val Ser Gly Pro Glu Leu
 180 185 190
 Gln Ala Ala Phe Pro Leu Asp Arg Leu Thr Pro Arg Val Ala Ala Trp
 195 200 205
 Asp Glu Ser Val Arg Ser Ala Leu Ala Leu Gly His Pro Ala Gly Phe
 210 215 220
 Tyr Pro Cys Pro Asp Ser Ala Phe Gly Leu Ser Arg Val Gly Val Met
 225 230 235 240
 His Phe Ala Ser Pro Ala Asp Pro Lys Val Phe Phe Arg Gln Thr Leu
 245 250 255
 Gln Gln Gly Glu Ala Leu Ala Trp Tyr Val Thr Gly Asp Ala Ile Leu
 260 265 270
 Asp Leu Thr Asp Arg Arg Ala Lys Thr Ser Pro Ser Arg Ala Met Gly
 275 280 285
 Phe Leu Val Asp Ala Ile Val Arg Val Ala Ile Asn Gly Trp Val Cys
 290 295 300
 Gly Thr Arg Leu His Thr Glu Gly Arg Gly Ser Glu Leu Asp Asp Arg

```

305      310      315      320
Ala Ala Glu Leu Arg Arg Gln Phe Ala Ser Leu Thr Ala Leu Arg Pro
      325      330      335
Val Gly Ala Ala Ala Val Pro Leu Leu Ser Ala Gly Gly Ala Ala Pro
      340      345      350
Pro His Pro Gly Pro Asp Ala Ala Val Phe Arg Ser Ser Leu Gly Ser
      355      360      365
Leu Leu Tyr Trp Pro Gly Val Arg Ala Leu Leu Gly Arg Asp Cys Arg
      370      375      380
Val Ala Ala Arg Tyr Ala Gly Arg Met Thr Tyr Ile Ala Thr Gly Ala
385      390      395      400
Leu Leu Ala Arg Phe Asn Pro Gly Ala Val Lys Cys Val Leu Pro Arg
      405      410      415
Glu Ala Ala Phe Ala Gly Arg Val Leu Asp Val Leu Ala Val Leu Ala
      420      425      430
Glu Gln Thr Val Gln Trp Leu Ser Val Val Val Gly Ala Arg Leu His
      435      440      445
Pro His Ser Ala His Pro Ala Phe Ala Asp Val Glu Gln Glu Ala Leu
      450      455      460
Phe Arg Ala Leu Pro Leu Gly Ser Pro Gly Val Val Ala Ala Glu His
465      470      475      480
Glu Ala Leu Gly Asp Thr Ala Ala Arg Arg Leu Leu Ala Thr Ser Gly
      485      490      495
Leu Asn Ala Val Leu Gly Ala Ala Val Tyr Ala Leu His Thr Ala Leu
      500      505      510
Ala Thr Val Thr Leu Lys Tyr Ala Leu Ala Cys Gly Asp Ala Arg Arg
      515      520      525
Arg Arg Asp Asp Ala Ala Ala Ala Arg Ala Val Leu Ala Thr Gly Leu
      530      535      540
Ile Leu Gln Arg Leu Leu Gly Leu Ala Asp Thr Val Val Ala Cys Val
545      550      555      560
Ala Leu Ala Ala Phe Asp Gly Gly Ser Thr Ala Pro Glu Val Gly Thr
      565      570      575
Tyr Thr Pro Leu Arg Tyr Ala Cys Val Leu Arg Ala Thr Gln Pro Leu
      580      585      590
Tyr Ala Arg Thr Thr Pro Ala Lys Phe Trp Ala Asp Val Arg Ala Ala
      595      600      605
Ala Glu His Val Asp Leu Arg Pro Ala Ser Ser Ala Pro Arg Ala Pro
      610      615      620
Val Ser Gly Thr Ala Asp Pro Ala Phe Leu Leu Glu Asp Leu Ala Ala
625      630      635      640
Phe Pro Pro Ala Pro Leu Asn Ser Glu Ser Val Leu Gly Pro Arg Val
      645      650      655
Arg Val Val Asp Ile Met Ala Gln Phe Arg Lys Leu Leu Met Gly Asp
      660      665      670
Glu Glu Thr Ala Ala Leu Arg Ala His Val Ser Gly Arg Arg Ala Thr
      675      680      685
Gly Leu Gly Gly Pro Pro Arg Pro
690      695

```

<210> 239

<211> 19

<212> PRT

<213> Herpes simplex virus

<400> 239

```

Leu Gly Arg Val Gly Gly Ser Arg Val Val Pro Ser Pro Leu Phe Leu
 1           5           10           15

```

Asp Glu Leu

<210> 240

<211> 725

<212> DNA

<213> Herpes simplex virus

<400> 240

```

ccacacaaaa cccccgccgg cgcgtctcca gaaacgcca caaccaaggg ggctcgccacc 60
ccgcgtcggc gcggacggac ggccccggcg ccaccacgg cgaggcgagg cgcgaggagg 120
agcagctcga cgtctccggg ggccccggcg cacgaggcac gcgccaggcc cccctccgc 180
tgatggcgct gtccctgacc cccccgcacg cggacggccg cggcccggtc ccggagcgaa 240
aggcgccctc tgccgacacc atcgaccccg ccgttcgggc gggtctgcga tccatatccg 300
agcgcgcgcg ggctcgagcg atcagcgaaa gctttggacg cagtgccttg gtcattgcaag 360
acccctttgg cgggatgcgg tttcccgccg cgaacagccc ctgggctccc gtgctggcca 420
cccaagcggg ggggtttgac gccgagaccc gtcgggtttc ctgggaaacc ctggctcgctc 480
acggcccgag cctctaccgc acattcgag ccaaccgcgg ggccgcgtcg acagccaagg 540
ccatgcgcga ctgcgtgctg cggcaggaaa atctcatcga ggccctggcg tccgcggatg 600
agacgtggc gtggtgcaag atgtgcattc accacaatct gccgctccgc cccaggacc 660
ctatcatcgg aacggcgggc gccgtgctgg aaaacctcgc cgcgcgctg cggcccttc 720
tgag 725

```

<210> 241

<211> 241

<212> PRT

<213> Herpes simplex virus

<400> 241

```

Thr Pro Asn Pro Arg Arg Arg Val Ser Arg Asn Ala His Asn Gln Gly
1 5 10 15
Gly Arg His Pro Ala Ser Ala Arg Thr Asp Gly Pro Gly Ala Thr His
20 25 30
Gly Glu Ala Arg Arg Gly Gly Glu Leu Asp Val Ser Gly Gly Pro
35 40 45
Arg Pro Arg Gly Thr Arg Gln Ala Pro Pro Pro Leu Met Ala Leu Ser
50 55 60
Leu Thr Pro Pro His Ala Asp Gly Arg Ala Pro Val Pro Glu Arg Lys
65 70 75 80
Ala Pro Ser Ala Asp Thr Ile Asp Pro Ala Val Arg Ala Val Leu Arg
85 90 95
Ser Ile Ser Glu Arg Ala Ala Val Glu Arg Ile Ser Glu Ser Phe Gly
100 105 110
Arg Ser Ala Leu Val Met Gln Asp Pro Phe Gly Gly Met Pro Phe Pro
115 120 125
Ala Ala Asn Ser Pro Trp Ala Pro Val Leu Ala Thr Gln Ala Gly Gly
130 135 140
Phe Asp Ala Glu Thr Arg Arg Val Ser Trp Glu Thr Leu Val Ala His
145 150 155 160
Gly Pro Ser Leu Tyr Arg Thr Phe Ala Ala Asn Pro Arg Ala Ala Ser
165 170 175
Thr Ala Lys Ala Met Arg Asp Cys Val Leu Arg Gln Glu Asn Leu Ile
180 185 190
Glu Ala Leu Ala Ser Ala Asp Glu Thr Leu Ala Trp Cys Lys Met Cys
195 200 205
Ile His His Asn Leu Pro Leu Arg Pro Gln Asp Pro Ile Ile Gly Thr
210 215 220
Ala Ala Ala Val Leu Glu Asn Leu Ala Thr Arg Leu Arg Pro Phe Leu

```

225
Gln

230

235

240

<210> 242
 <211> 1539
 <212> DNA
 <213> Herpes simplex virus

<400> 242
 atggctaccg acattgatat gctaatacgac ctaggattgg acctgtccga cagcgagctc 60
 gaggaggacg ctctggagcg ggacgaggag ggccgcgcgc acgaccccg gtccgacacg 120
 agcggggaggt gttcctcgtc ggacgaggac atggaagacc cctgcggaga cggaggggag 180
 gaggccatcg acgcggcgat tcccaaaggt ccccgggccc gcccggagga cgcgggcacc 240
 cccgaagcct cgacgcctcg cccggcagcg cggcggggag cgcacgatcc gccaccgcgc 300
 accaccggcg tgtgtgcgcg cctcgggacc aggcggtcgg cttccccccg ggaaccgcac 360
 ggggggaagg tggcccgcat ccaacccccg tcgaccaagg caccgcatcc ccgaggcggg 420
 cggcgaggtc gccgccgggg ccggggtcga tacggccccg gcggcgccga ctccacacca 480
 aaaccccgcc ggcggtctc cagaaacgcc cacaaccaag ggggtcgcca ccccggtcg 540
 gcgcggacgg acggccccg gcacacccac ggcgaggcgc ggcgcgagg ggagcagctc 600
 gacgtctccg ggggcccgc gccacgaggc acgcgccagg cccccctcc gctgatggcg 660
 ctgtccctga ccccccgca cgcggacggc cgcgcccgcg tcccgagcg aaaggcgccc 720
 tctgccgaca ccatcgaccc cgcggttcgg gcggttctgc gatccatata cgagcgcgcg 780
 gcggtcgagc gcatcagcga aagctttgga cgcagtgcc tggatcatgca agaccccttt 840
 ggcgggatgc cgtttccgc cggaacacgc ccctgggctc ccgtgctggc caccgaagcg 900
 gggggggttg acgcccagac ccgtcggtt tcctgggaaa ccctggtcgc tcacggccccg 960
 agcctctacc gcacattcgc agccaacccg cggcccgct cgacagccaa ggccatgcgc 1020
 gactgcgtgc tgcgccagga aaatctcatc gaggccctgg cgtccgcgga tgagacgctg 1080
 gcgtggtgca agatgtgcat taccacaaat ctgccgtcc gccccagga ccctatcatc 1140
 ggaacggcgg ccgcgtgct ggaaaacctc gccacgcgcc tgcgccctt tctgcagtgc 1200
 tacctgaagg cccgaggcct gtgcgggctg gacgacctgt gctcgcggcg acgcctgtcg 1260
 gacattaagg atattgcctc ctttgtgttg gtcaccttg cccgcctcgc caaccgcgtc 1320
 gagcgcgcg tgtcggagat cgactacacg accgtggggg ttggggcccg cgagacgatg 1380
 cacttttaca tcccgggggc ctgcatggcg ggtctcattg aaatactgga cagcaccgc 1440
 caggagtgtt ccagtcgcgt gtgcgagctg acggccagtc aactatcgc ccccttatat 1500
 gtgcacggca aatacttcta ctgcaactcc ctattttag 1539

<210> 243
 <211> 512
 <212> PRT
 <213> Herpes simplex virus

<400> 243
 Met Ala Thr Asp Ile Asp Met Leu Ile Asp Leu Gly Leu Asp Leu Ser
 1 5 10 15
 Asp Ser Glu Leu Glu Glu Asp Ala Leu Glu Arg Asp Glu Glu Gly Arg
 20 25 30
 Arg Asp Asp Pro Glu Ser Asp Ser Ser Gly Glu Cys Ser Ser Ser Asp
 35 40 45
 Glu Asp Met Glu Asp Pro Cys Gly Asp Gly Gly Ala Glu Ala Ile Asp
 50 55 60
 Ala Ala Ile Pro Lys Gly Pro Pro Ala Arg Pro Glu Asp Ala Gly Thr
 65 70 75 80
 Pro Glu Ala Ser Thr Pro Arg Pro Ala Ala Arg Arg Gly Ala Asp Asp
 85 90 95
 Pro Pro Pro Ala Thr Thr Gly Val Trp Ser Arg Leu Gly Thr Arg Arg
 100 105 110
 Ser Ala Ser Pro Arg Glu Pro His Gly Gly Lys Val Ala Arg Ile Gln

115	120	125
Pro Pro Ser Thr Lys Ala	Pro His Pro Arg Gly Gly	Gly Arg Arg Gly Arg
130	135	140
Arg Arg Gly Arg Gly Arg	Tyr Gly Pro Gly Gly	Ala Asp Ser Thr Pro
145	150	155
Lys Pro Arg Arg Arg Val	Ser Arg Asn Ala His	Asn Gln Gly Gly Arg
165	170	175
His Pro Ala Ser Ala Arg	Thr Asp Gly Pro Gly	Ala Thr His Gly Glu
180	185	190
Ala Arg Arg Gly Gly Glu	Gln Leu Asp Val Ser	Gly Gly Pro Arg Pro
195	200	205
Arg Gly Thr Arg Gln Ala	Pro Pro Leu Met Ala	Leu Ser Leu Thr
210	215	220
Pro Pro His Ala Asp Gly	Arg Ala Pro Val Pro	Glu Arg Lys Ala Pro
225	230	235
Ser Ala Asp Thr Ile Asp	Pro Ala Val Arg Ala	Val Leu Arg Ser Ile
245	250	255
Ser Glu Arg Ala Ala Val	Glu Arg Ile Ser Glu	Ser Phe Gly Arg Ser
260	265	270
Ala Leu Val Met Gln Asp	Pro Phe Gly Gly Met	Pro Phe Pro Ala Ala
275	280	285
Asn Ser Pro Trp Ala Pro	Val Leu Ala Thr Gln	Ala Gly Gly Phe Asp
290	295	300
Ala Glu Thr Arg Arg Val	Ser Trp Glu Thr Leu	Val Ala His Gly Pro
305	310	315
Ser Leu Tyr Arg Thr Phe	Ala Ala Asn Pro Arg	Ala Ala Ser Thr Ala
325	330	335
Lys Ala Met Arg Asp Cys	Val Leu Arg Gln Glu	Asn Leu Ile Glu Ala
340	345	350
Leu Ala Ser Ala Asp Glu	Thr Leu Ala Trp Cys	Lys Met Cys Ile His
355	360	365
His Asn Leu Pro Leu Arg	Pro Gln Asp Pro Ile	Ile Gly Thr Ala Ala
370	375	380
Ala Val Leu Glu Asn Leu	Ala Thr Arg Leu Arg	Pro Phe Leu Gln Cys
385	390	395
Tyr Leu Lys Ala Arg Gly	Leu Cys Gly Leu Asp	Asp Leu Cys Ser Arg
405	410	415
Arg Arg Leu Ser Asp Ile	Lys Asp Ile Ala Ser	Phe Val Leu Val Ile
420	425	430
Leu Ala Arg Leu Ala Asn	Arg Val Glu Arg Gly	Val Ser Glu Ile Asp
435	440	445
Tyr Thr Thr Val Gly Val	Gly Ala Gly Glu Thr	Met His Phe Tyr Ile
450	455	460
Pro Gly Ala Cys Met Ala	Gly Leu Ile Glu Ile	Leu Asp Thr His Arg
465	470	475
Gln Glu Cys Ser Ser Arg	Val Cys Glu Leu Thr	Ala Ser His Thr Ile
485	490	495
Ala Pro Leu Tyr Val His	Gly Lys Tyr Phe Tyr	Cys Asn Ser Leu Phe
500	505	510

<210> 244
 <211> 1921
 <212> DNA
 <213> Herpes simplex virus

<400> 244
 accgcggccc tggcgggcat cctcaggggtg gtccgaacgt gcgtcccatt tttgtggaga
 cgcaataaa cgggtgtgtg gacgcagcgg cgtccacgcc caaccaacc gactccctcc

60
 120

gtgtccgcgg	tctgtttgtt	attgtgtccg	ccgtgcgtcc	gctaccgcct	ctgttccttt	180
cccttctcca	ttcctgtttc	ctttcctttc	cccccccca	tagtcccccg	tataggcata	240
caacggcatc	cgtgggttag	aaaacgactg	cactttattg	ggatatctca	cacagactgg	300
ccgtgctggg	cgcgagccag	gcaaacggtg	agcagcggt	ccaggtacc	ggcggttcgc	360
gtgcggccag	tgggtccgac	gaggccttcc	cgatcgagta	cgctcctgcg	ctcatgaacg	420
actgggcca	tgtgccctgc	aaccctacc	tgcgcgtgca	gaacaccggc	gtttcgggtg	480
tgtttcagg	tttttttaac	cgccccacg	gcgccccggg	ggcgcgatc	acggcgagc	540
agaccaacgt	gattctgcac	tccaccgaga	cgacgggact	gtccctcgga	gacctggacg	600
acgtcaaggg	gcgcctcggc	ctggacgccc	ggcgatgat	ggccagcatg	tggatcagct	660
gctttgtgcg	catgcccccg	gtgcagctcg	cgtttcgggt	catgggcccc	gaggacgcgc	720
ttcgacgcg	gcggatcctg	tgtcgcgccg	ccgagcaggg	cctcgccccg	cgccgcgggt	780
ccaggcggtc	ccaggatgac	tacggggcgg	tggcggtggc	ggcgcgcac	cactcttccg	840
gagcgcccg	gccgggggtc	gccgcctcgg	gcccgcacg	gccgcccga	cggggacggg	900
ccgtccgtg	gcatcaggcc	gtgcagtgtg	tccgggcccc	gcgtccgggc	ccccggcgc	960
ttctgttgct	ggcgcgggg	ttgtttctgg	gggcgcctat	ctggtgggcg	gttggcgcg	1020
gcctatgaaa	ggggcgagc	caccgtccc	cccgccagtg	catcccagac	gcccgcgagc	1080
cgcacatccc	ctccgctccc	gcctccggcc	cgattcttac	ggcgcgacc	aagggtccga	1140
tggccgcccc	gcagtttcac	cgccccagca	ccattaccgc	cgacaacgtc	cgggcgctcg	1200
gcatgcgcgg	gctcgtgttg	gccaccaaca	acgctcagtt	catcatggat	aacagctacc	1260
cgcacccgca	cggaacgcag	ggtgcggtgc	gagagtttct	tcgcgggcag	gccgcggcgc	1320
tgacggacct	cggggtgacc	cacgccaaca	acacgttcgc	ccgcagcct	atgttcgcgg	1380
gcgacgcgcg	ggccgaatgg	ctgcggccct	cgttcgggtc	taagcgcacg	tattccccct	1440
ttgtcgttcg	cgaccccaag	acccccagca	ccccgtgagt	cctcgcgggg	tccttcgcgc	1500
gccgtctctc	gttgcccccc	ctttccccct	tccgggtggg	ttcaataaaa	aacaccaaca	1560
tacgatattc	gcgtttgata	cgtttattgg	ggggtgtagg	gccaacgat	cggcgattaa	1620
caacacaaaa	caatcgagcg	cgtctaacc	agtaacatgc	gcacgtgatg	taggctggtc	1680
agcacggcgt	tgctgcgctg	aaacagcgcc	ctgcgggtcc	gctgcagctg	ttgttgatg	1740
cgggcgcatg	cgcgatcaa	aaccgccagg	gcgctacgac	cggtgcttcg	tacgtagcgt	1800
cgcgacaaga	cggcatttgc	ctgtacgggc	aaggggccaa	attgcgagtg	tggtagctgg	1860
aggtggtcgg	cggccaatgg	gccgggtggt	tcgtcggcgg	ggggcaagtg	cggttcgggt	1920
c						1921

<210> 245

<211> 1599

<212> DNA

<213> Herpes simplex virus

<400> 245

atggagctca	gctatgccac	caccctgcac	caccgggacg	ttgtgtttta	cgtcacggca	60
gacagaaacc	gcgcctactt	tgtgtgcggg	gggtccggtt	attccgtagg	gcggcctcgg	120
gattctcagc	cgggggaaat	tgccaagt	ggcctggtgg	tccgggggac	aggcccaaaa	180
gaccgcatgg	tcgccaacta	cgtaacgaag	gagctccgcc	agcgcgccct	gcgggacgtg	240
cgccccgtgg	gggaggacga	ggtgttcctg	gacagcgtgt	gtctgctaaa	cccgaacgtg	300
agctccgagc	gagacgtgat	taataccaac	gacgttgaag	tgctggacga	atgcctggcc	360
gaatactgca	cctcgtgcg	aaccagcccg	gggtgctgg	tgaccggggg	gcgcgtgcgc	420
gcgcgagaca	gggtcatcga	gctatttgag	caccggcgca	tcgtcaacat	ttcctcgcgc	480
ttcgcgtaca	ccccctcccc	ctacgtatc	gccctggccc	aggcgccact	ccccggctc	540
ccgagctcgc	tggagcccc	ggtgagcgcc	ctgttgacg	gcattcccgc	cccgcccg	600
cccctggacg	cccgcgaccg	gcgcacggat	gtcgtgatca	cgggcaccg	cgccccaga	660
ccgatggccg	ggaccggggc	cgggggcgcg	ggggccaagc	gggccaccgt	cagcgagttc	720
gtgcaagtga	agcacatcga	ccgtgttgtg	tccccgagcg	tctcttcgcg	cccccccg	780
agcgcccccg	acgcgagctc	gccgcccccg	gggtccagg	aggccgcccc	gccgggcccc	840
ccgctcagg	agctgtgtgt	ggtgtttctac	gccggcgacc	gggcgctgga	ggagccccac	900
gccgagtcgg	gattgacgcg	cgaggaggtc	cgcgcctg	atgggttccg	ggagcaggcg	960
tggaaagctgt	ttgggtcgg	gggggtccg	cgggcggttc	tcggggcgcg	gctggccctg	1020
agccccagcc	aaaagctcgc	cgtctactac	tatctcatcc	accgggagcg	gcgcatgtcc	1080
cccttccccg	cgctcgtgcg	gctcgtcgg	cggtacatcc	agcgccacgg	cctgtacgtt	1140
cccgcgcccc	acgaaccgac	gttggccgat	gccatgaacg	ggctgttccg	cgacgcgctg	1200

gcggccggga	ccgtggccga	gcagctcctc	atgttcgacc	tcctcccgc	caaggacgtg	1260
ccggtgggga	gcgacgcgcg	ggccgacagc	gccgccctgc	tgcgctttgt	ggactcgcga	1320
cgcctgaccc	cgggggggtc	cgtctcgccc	gagcacgtca	tgtacctcgg	cgcgttcctg	1380
ggcgtgttgt	acgcgcggca	cggacgcctg	gccgcggcca	cgcataccgc	gcgcctgacg	1440
ggcgtgacgt	ccctggtcct	gaccgtgggg	gacgtcgacc	ggatgtccgc	gtttgaccgc	1500
gggcccggcg	gggcggctgg	ccgcaecgca	accgcgggtg	acctggacgc	gctgcttacc	1560
gtttgcctgg	ctcgcgcccc	gcacggccag	tctgtgtga			1599

<210> 246

<211> 2517

<212> DNA

<213> Herpe simplex virus

<400> 246

atgggccccg	gtctgtgggt	ggtgatgggg	gtcctggtgg	gcgttgccgg	gggccatgac	60
acgtactgga	cggagcaaat	cgaccctggt	tttttgcaag	gtctgggggt	ggcccgacag	120
tactggcgcg	acacaaacac	cgggcgtctg	tggttgcccc	acacccccga	cgccagcgac	180
ccccagcgcg	gacgcttggc	gccccggggc	gaactcaacc	tgactacggc	atccgtgccc	240
atgcttcggt	ggtacgcgga	gcgcttttgt	ttcgtgttgg	tcaccacggc	cgagtttctt	300
cgggaccccc	ggcagctgct	ttacatccca	aagacctatc	tgtcggcccg	gcctcggaac	360
gcgagcctgc	ccgagctccc	cgaggcgggg	cccacgtccc	gtccccccgc	cgaggtgacc	420
cagctcaagg	gactgtcgca	caacccccgc	gcctccgcgc	tgttgcggtc	ccgggcctgg	480
gtaacattcg	cggccgcgcc	ggaccgcgag	gggcttacgt	tcccgcgggg	agacgacggg	540
gcgaccgaga	ggcaccggga	cggccgacgc	aacgcgcgcg	ccccggggcc	gcccgcgggg	600
acaccgaggc	atccgacgac	gaacctgagc	atcgcgcata	tgcacaacgc	atccgtgacc	660
tggctggcgc	ccaggggcct	gctacggact	ccgggtcggt	acgtgtacct	ctcccgtcgc	720
gcctcgacgt	ggcccgtggg	gctctggacg	acgggcgggc	tggcgttcgg	gtgcgacgcc	780
gcgctcgtgc	gcgcgcgata	cgggaagggc	ttcatggggc	tcgtgatata	gatgcgggac	840
agccctccgg	ccgagatcat	agtgggtgct	gcggacaaga	ccctcgctcg	ggtcggaaat	900
ccgaccgacg	aaaacgcccc	cgcggtgtct	cccgggcctc	cggccggccc	caggtatcgc	960
gtctttgtcc	tggggggccc	gacgcccgcg	gacaacggct	cggcgctgga	cgccctccgg	1020
cgggtggccg	gctaccccga	ggagagcacg	aactacgccc	agtatatgtc	gcgggcctat	1080
gcggagtgtt	tgggggagga	cccgggctcc	ggcaocggacg	cgcgtccgtc	cctgttctgg	1140
cgctcgcgg	ggctgctcgc	ctcgtcgggg	tttgcggttcg	tcaacgcggc	ccacgcccac	1200
gacgcgattc	gcctctccga	cctgctgggc	tttttgggcc	actcgcgcgt	gctggccggc	1260
ctggccgccc	ggggagcagc	gggctgcgcg	gccgactcgg	tgttcctgaa	cgtgtccgtg	1320
ttggaccggg	cggcccgcct	gcggctggag	gcgcgcctcg	ggcatctggt	ggccgcgata	1380
ctcgaagcag	agcagagcct	ggtggcgcac	gcgctgggct	atcagctggc	gttcgtgttg	1440
gacagccccg	cggcctatgg	cgcggtggcc	ccgagcgcgg	cccgcctgat	cgacgccttg	1500
tacgccgagt	ttctcggcgg	ccgcgcgcta	accgcccga	tggcccgccg	agcgtgtgtt	1560
tacgccacgg	ccgtcctccg	ggcgccgttc	ctggcggggc	cgcctccggc	cgagcagcgg	1620
gaacgcgccc	gcccggggcct	cctcataaac	acggccctgt	gtacgtccga	cgtcgcgcgc	1680
gcgacccacg	ccgatctccg	ggccgcgcta	gccaggaccg	accaccagaa	aaacctcttc	1740
tggctcccgg	accacttttc	cccatgcgca	gcttccctgc	gcttcgatct	cgccgagggc	1800
gggttcatcc	tggacgcgct	ggccatggcc	acccgatccg	acatcccggc	ggacgtcatg	1860
gcacaacaga	cccgcggcgt	ggcctccggt	ctcacgcgct	ggcgcgacta	caacgccttg	1920
atccgcgcct	tcgtcccggg	ggccaccac	cagtgtagcg	gcccgtcgca	caacgcggag	1980
ccccggatcc	tcgtgcccat	cacccacaa	gccagctacg	tcgtcaccca	cacccctttg	2040
ccccgcggga	tcggatacaa	gcttacgggc	gttgacgtcc	gccgcccgtc	gtttatcacc	2100
tatctcaacc	ccacctgcga	agggcacgcg	cgggagattg	agccgaagcg	gctggtgcgc	2160
accgaaaacc	ggcgcgacct	cggcctcggt	ggggccgtgt	ttctgcgcta	caccccgccc	2220
ggggaggtca	tgtcgggtgt	gctggtggac	acggatgcca	cccaacagca	gctggcccag	2280
gggcccgttg	cgggcacccc	gaacgtgttt	tcacgcgacg	tgcggtccgt	ggccctgttg	2340
ttgttcccca	acggaactgt	gattcatctg	ctggcctttg	acacgtcgcc	catcgccacc	2400
atcgcccccg	ggtttctggc	cgcgtccgcg	ctgggggtcg	ttatgattac	cgcggccctg	2460
gcgggcatcc	ttaggggtgt	ccgaacgtgc	gtcccatttt	tgtggagacg	cgaataa	2517

<210> 247

<211> 9369

<212> DNA

<213> Herpes simplex virus

<400> 247

atgattccgg	ccgcgctacc	tcacccacg	atgaagcgcc	agggggatcg	cgatatcgtg	60
gtcaccggtg	ttcggaaacca	gttcgcgacc	gatctggagc	cgggggggtc	ggtatcgtgc	120
atgcgctcgt	ccctgtcgtt	tctcagtcct	cttttcgacg	tggggccccc	cgacgtgctg	180
tccgcggagg	ccatcgaggg	atgtctgggtg	gagggggggc	agtggacgcg	cgcgcccgcg	240
ggctccgggc	cgccgcgcat	gtgttccatt	atcgagctcc	ccaacttcct	ggagtacccg	300
gcggcgcgcg	gcggcctgcg	ttgcgtgttt	tcgcgggtgt	acggcgaggt	gggtttcttc	360
ggggagccca	cggcggggct	gctggagacg	caatgccccg	cgcacacgtt	cttcgcgggc	420
ccctgggcca	tgcggccgct	gtcctacacg	ctcctgacca	tcgcccccct	gggcatgggg	480
ctctaccggg	acggcgacac	ggcctatctg	tttgaccgcg	acggcctgcc	cgcggaacg	540
cccgcgttta	tgcgaaggt	gcgggcgggc	gacgtgtacc	cgtacctgac	gtattacgcc	600
cacgaccgcc	ccaaggtgcg	gtgggcgggc	gccatggtgt	tttttgtgcc	ctccggcccg	660
ggggccgtgg	cgccggcgga	cctgaccgcc	gcggccctgc	acctctacgg	ggccagcgag	720
acttacctgc	aggacgaacc	gttcgtcgag	cggcgcgtgg	ccatcaccca	ccccctgcgg	780
ggcgagatcg	gggggtcggg	ggcgctcttt	gtcgggtgtg	tcccagggg	ggacggagaa	840
ggctcggggc	cggtcgtccc	cgcgctcccc	gcgcccaccc	acgtccagac	ccctggggcc	900
gaccggcccc	ccgaggcccc	ccgggggggc	tcggggccac	ccgacacacc	gcaggcgggg	960
cacccaaacc	gcccccccca	tgatgtgtgg	gcagcagccc	tggaggggac	cccggcccg	1020
aagccctcgg	ccccggacgc	cgcggcctcg	ggtccaccgc	acgcgcgcc	gcccccgag	1080
acaccgccc	gggacgcggc	cgaagaagcc	gaggacctgc	gggtcctgga	ggtgggggca	1140
gtgcgcgtgg	gtcgccaccg	ggcgcgctat	tctaccggct	tgcccaagcg	cgccggccc	1200
acctggaccc	caccgtccag	cgtggaggat	ctgacctcgg	gcgagaggcc	cgccccaaag	1260
gccccgcccc	ccaaggccaa	gaagaagtcc	gcgccaaga	agaaggcccc	cgtcgccgcc	1320
gaggtcccc	cctcctcccc	gacgcccata	gccgccacgg	tccctccgcg	accggatacc	1380
cccccgcagt	ccgggcagg	cggcggcgac	gacgggcccc	cgagcccgtc	gtcgccgtcc	1440
gtgctcgaga	cgctgggggc	ccgcgccccc	ccggagcccc	ccggcgccga	cctggcgag	1500
ctgttcgagg	tccaccccaa	tgtggccgcc	acggcggtcc	ggctcgcggc	gcgcgatgcc	1560
gccctggccc	gcgaggtggc	cgcggtttcg	cagctcacca	tcaacgcctt	gcggtcgccg	1620
tacccggcgc	acccggggct	gttgagctc	tgcgtcatct	tcttctttga	gcgcgtcctc	1680
gcctttttga	tcgagaacgg	ggcccgcacg	catacgagc	ccggggtggc	cgccccggcc	1740
cgggccctgc	tggacttcac	cctgcgcagt	ctgcccgcga	aaacggccgt	cggggacttc	1800
ctggcgctcca	cgcgcatgag	cctggcggag	gtggccgcgc	accggccctt	catccagcac	1860
gtgctggacg	aaaattcgca	gatcgggcgc	ctggccctcg	ctaagctggt	cctggtcgcc	1920
agggacgtta	ttcgggagac	ggacgccttt	tacggggacc	tcgcggaact	ggatctgcag	1980
ctgcgcgcgc	ccccgcgggc	caatctgtac	gcacgcctcg	gagagtggct	cctggagcgc	2040
tcgcgcgcgc	accogaacac	cctgttcgcc	cccgcacccc	cgacgcaccc	ggagccgctt	2100
ctccatcgga	tccaggcgct	ggcccagttt	gcccggggcg	aagagatgcg	ggtcgaggcg	2160
gaggcccgcg	agatgcgcga	ggccttagac	gccctcgccc	gcggggtcga	ctcggtctcg	2220
cagcgcgccg	ggccctgac	cgtgatgccc	gtgcccgcgc	ccccgggcgc	cggcggcagg	2280
gccccgtgcc	cgcccgccct	gggccccgag	gccatccagg	cgcggtgga	ggacgtgcgg	2340
atccaggccc	gcggggcgat	cgagagcgcg	gtcaaggagt	acttcacccg	gggagccgta	2400
tacagcgcca	aggccctgca	ggccagcgag	agccacgact	gtcggtttca	cgtggcctcg	2460
gccgcggtcg	tgccattggt	ccagttgctg	gaatcgctac	cggcctttga	ccaacacacg	2520
cgggacgtcg	ctcagcgcg	cgcgctcccc	ccccgcgcc	ccctcgccac	cagccccag	2580
gccatcctgc	tgcgggacct	gctccagcgg	ggccagcccc	tagatgcccc	cgaggacctg	2640
gcggcctggc	tctcggtcct	gaccgacgcc	gcgacccagg	ggctcatcga	acgcaagccc	2700
ctggaagagc	tggcgcgag	catccacggc	atcaacgacc	agcaggcgcg	gcgcagctcg	2760
gggctggccg	aactgcagcg	cttcgacgcc	ctcgacgcgg	cgctcgcgca	gcagctggac	2820
agcgacgcgg	ccttcgtccc	ggcgaccggc	ccgcgcgcgt	acgtcgacgg	gggcgggctg	2880
tcccccgagg	ccacgcgcat	ggccgaagac	gcgctgcggc	aggcgagggc	catggaggcg	2940
gccaagatga	cggcggagct	cgctccggag	gcgcgctccc	ggcttcggga	gcgcgcccac	3000
gccctggagg	cgatgctcaa	cgacgcgcgg	gagcgcgcca	aggtggccca	cgacgcccgg	3060
gagaagtctc	tgcacaaact	ccaagggtgc	ctgcggcccc	tcccggactt	tgtggggcta	3120
aaggcgtgtc	cggccgtgct	ggcgaccctg	cgggcctccc	tgcccgcggg	gtggaccgac	3180

ctggccgatg	ccgtccgggg	ccccccccc	gaagtcacgg	cggcgctgcg	ggcggacctg	3240
tgggggctgc	tggggcagta	ccgggaggcc	ctggaacacc	cgaccccggg	taccgcgacg	3300
gctctggccg	gcctgcatcc	cgccttcgtg	gtggtgctga	agaccctgtt	cgccgacgcc	3360
cccagagacg	cggtcctcgt	gcagttcttt	tccgatcacg	ccccacaaat	cgccaaagcc	3420
gtgtcaaacg	ccatcaacgc	cgggagcgcc	gccgtcgcga	cggcaagtcc	cgcggcgacg	3480
gtggacgcgg	ccgtgcgggc	gcatggcgcc	ctggcggaag	cgggtgtcggc	cctggggcgcg	3540
gccgcccgcg	acccggcctc	gcccctgtcc	ttcctcgcgg	tgtctggccga	tagcgccgcg	3600
gggtatgtaa	aggcgactcg	gctggccctg	gagggcgggg	gggccatcga	cgagctcacg	3660
accctggggt	cggcgggccg	cgacctggtg	gtccaggcgc	gccgggcctg	cgcccagccg	3720
gagggggacc	atgccgccct	catcgatgcc	gcggcccgcg	cgacgaccgc	tgctcgggaa	3780
agcctcgccg	ggcacgaggc	tgggttcggg	ggcctgctgc	acgcggaagg	gacggccggg	3840
gaccattccc	ccagcggggc	cgcctgcag	gagctgggca	aggtcatcgg	cgccacgcga	3900
cgcgcgcggc	acgaacttga	ggccgcccgt	gccgacctga	cggcgaagat	ggcggcccag	3960
cgcgcgggtca	ccagcagcga	gcgtggggcc	gccggcgctg	aggcggcgct	ggaccgcgtc	4020
gagaaccgcg	ccgagttcga	cgtggtcgag	ctgcgcgcgc	tgcaggcgct	ggcgggcacg	4080
cacggctaca	acccccggga	cttcgcgaag	cgcgcgcgag	aggcgctcgc	ggccaacgcc	4140
gagggcggtca	ccctcgccct	ggatacggcc	ttcgcgttta	acccctacac	ccccgagaac	4200
cagcgccacc	ccatgctccc	cccgtcgcgc	gccatccacc	gcctcggtcg	gagcgcggcg	4260
ttccacgcgc	ccgcccgaac	ctacgcccgc	atgtttcggg	tcgacgccga	gcccctggcg	4320
cggctgctgc	gcatcgcgga	ggggctcctg	gaaatggccc	aggcggggcg	cgggtttatc	4380
gactaccacg	aggcgtggg	gcgcctggcg	gacgacatga	cgagcgtgcc	cggcctgcgc	4440
cggtaoctgc	cgttcttcca	acatggctac	gccgactacg	tggagctccg	cgatcgccctg	4500
gacgccatcc	gggcgcagct	gcaccggggc	ctcggcgggc	tgcccctgga	cctggcggcc	4560
gccgcgggag	agatatccgc	ggcccgggat	gaccgcggag	ccaccgcgga	gctggtccgg	4620
accgggggtca	ccctgcgcgtg	tcccagcgag	gacgcgctgg	tcgcgtgcgc	ggcggcggtg	4680
gagcgcgtgg	accagagccc	cgtgaagaac	acggcgatcg	ccgagtatgt	cgccttcgtg	4740
acccggcagg	acacggccga	gaccaaggac	gcagtgggtg	gcgccaaaca	gcagcgcgcg	4800
gagggccaccg	agcgggtcat	ggcgggcctg	cgggaggcgc	tggccgcgcg	cgagcgccgg	4860
gcccagatcg	aggccgaggg	cctggccaac	ctcaagacca	tgctgaaggt	ggtcgcgcgtc	4920
ccggcgaccg	tggccaagac	gctcgaccag	gcgcgctcgg	tggcggagat	cgccgaccag	4980
gtcgaggttc	tggtggatca	gaccgagaaa	acgcgcgagc	tcgacgtgcc	cgcggttatc	5040
tggetagagc	acgcccagcg	caccttcgag	acgcacccgc	taagcgccgc	ccgcggcgac	5100
ggccccggcc	ccctggcgcg	gcacgcgcgc	cgcctggggg	cgctctttga	tactcggcgc	5160
cgcgtcgacg	ccctgcggag	gtccctggag	gagggcggag	cggagtggga	cgaggtctgg	5220
ggtcgctttg	gccgcgtccg	ggggggcgcc	tggaaagtccc	ccgaggggtt	ccgcgccatg	5280
cacggaacgc	tgcgcgccct	gcaggacacc	accaaaccgg	tgtcggggct	gcggggcccag	5340
ccggcctacg	agcgcctttc	cgcaggtac	caggcgctcc	tgggcgccaa	gggcgcgcag	5400
cgcgcggagg	ccgtggagga	gctggggcgcg	cgcgtcacca	agcacaccgc	cctgtgcgcgc	5460
cggctgcggg	atgaggtggt	gcgcagggtg	ccctgggaga	tgaactttga	cgcctggggg	5520
ggcctgctgg	cggagttcga	cgcggcgggc	gcggacctgg	ccccgtgggc	ggtggaggag	5580
ttccggggcg	cgcgggaact	catccaatac	cgcattgggt	tatatagcgc	gtacgccagg	5640
gccggcgggc	agacgggcgc	gggcgcggag	tcgcgcgccg	cgcctctgct	cgttgatctt	5700
cgcgcctcgc	acgcccgcgc	ccgggcctct	tcttcccccg	agggccacga	ggtcgacccg	5760
cagctgctgc	gccgcggggg	cgaggcgtat	ctgcgcgcgc	gcggagaccc	ggggcccctg	5820
gtgctgcgtg	aggccgtcag	cgtctggat	ctgcgcttcg	ccaccagett	tttggccccg	5880
gacggaacgc	cgtgcagta	cgcgtatgc	ttccggccgc	tcaccgacaa	actcggcgcg	5940
ctgctgatgc	gtcccagggc	ggcctgcgtg	cggccccgcg	ttccgacgga	cgtcctcgaa	6000
tgggccccga	cgttcacggc	catgtacgtg	ctgacccctg	tgaaccgggt	ccagctggcc	6060
ctcagcgacg	cccaggccgc	caactttcag	ctgttcggtc	gcttcgtgcg	ccatcgccag	6120
gcgacgtggg	gcgcctcgat	ggacgcggcg	gccgagctgt	acgtcgccct	cgtcgccacc	6180
accctcacgc	gcgagtttgg	gtgtcgctgg	gccagctgg	gctgggcgtc	cggagcgggc	6240
gcgcgcgcgc	cgcgcggggg	ccccggggg	tcccagcgcc	actgcgtcgc	cttcaacgag	6300
aacgacgtgc	tggtcgcgct	ggtggccggc	gttcgggaac	acatctacaa	cttctggcgc	6360
ctggacctcg	ttcgccagca	cgagtacatg	cacctacccc	tcgaacgcgc	gttcgaggac	6420
gcagcggagt	ccatgctgtt	cgtccagcgc	ctgacccgcg	atcccagcgc	ccgcatccgc	6480
gtgttgcgga	cgttttttga	cggaggcccc	ccgacccggg	gcctcctgtt	cggcacgcgg	6540
ctggccgact	ggcgccgggg	caagctgtcc	gaaaccgacc	cgtggcgccc	ctggcgctcg	6600
gccttgagag	tggggaccca	gcgcggggac	gtcccgcgcg	tggggaagct	cagtcgggcc	6660

caggccctgg	cggcggtgag	cgccctcggg	cgcatgtgtc	tgccgagcgc	cgctttggcc	6720
gcgctgtgga	cctgcatgtt	tcccagcgac	tacaccgagt	acgacagctt	cgacgccttc	6780
ctggccgcac	gcctggagtc	tggccagacg	ctcggeccgg	cgggggggcg	cgaggcggtcc	6840
ctccccgagg	ccccccacgc	cctctaccga	cccacgggcc	agcacgtggc	cgtgctggcc	6900
gccgcgaccc	accgcacccc	cgccgcgcgc	gttacggcca	tggacctggt	tctggccgcg	6960
gtgctcctcg	gcgcgcccgt	cgtggtggcg	ctccgcaaca	ccacggcctt	ctccgcgag	7020
tcggaactgg	aactgtgcct	gacgtctctc	gactcgcgcc	ccggcggggc	ggacgcgcgc	7080
ctgcgcgacg	tcgtgtcgtc	cgacatcgag	acgtggggcg	tcggcctcct	ccacaccgat	7140
ctcaaccgga	tcgaaaacgc	gtgtctggcg	gcgcagctcc	cgcgcctgtc	ggcgctcatc	7200
gccgagcgcc	ctctcgccga	cgggcccccg	tgccctggtc	tcgtggacat	ctccatgacc	7260
ccggtcgcgg	tcttgtggga	agccccggag	ccccccggcc	cccctgacgt	gcggtttgtg	7320
ggcagcgagg	ccaccgagga	gcttcggttt	gtggctaccg	cgggggacgt	tctcgcggcg	7380
agcgccggcg	acgcggaccc	cttcttcgcg	cgggccatcc	tcgggcggcc	cttcgacgcc	7440
tccctcctga	cgggggagct	gttcccggga	caccgggttt	accagcgccc	cctcgcggac	7500
gaggcaggtc	cctctgcccc	gaccgcgcgc	cgcgacccgc	gggaccttgc	ggggggggat	7560
ggcgatcg	gtcccagga	ccccgctgcc	ccccccgcac	ggcaggcgga	cccgggggtc	7620
ctcgccccca	ctctcctcac	cgacgccacc	accggcgagc	ccgtcccccc	tcgcatgtgg	7680
gcctggatcc	acggcctgga	ggagctggcg	tccgacgacg	ccggcggccc	cacgcccatt	7740
ccggcccccg	cccttctctc	cccccccgcc	accgatcagt	ccgtccccac	gtcccagtac	7800
gcaccgcggc	ccatcgggcc	ggcagctacg	gtcgcgaaa	caagaccgag	tgtcccgcct	7860
caacaaaaca	cggggcgcgt	gcccgtggcc	cctcgcgacg	accacggccc	ctcgccaccc	7920
acaccgagtc	ccccgcggga	tgccgcgctt	cctcccccg	ccttttccgg	ctctgccgcc	7980
gctttctccg	ccgcgctgcc	gcgcgtgcgc	aggtcccgc	ggacgcgtgc	caaatcgcg	8040
gcgccccggg	cctcgcccc	ccccgagggg	tggcgtccgc	cggcactccc	cgccccggtc	8100
gccccggctg	ccgcttccgc	gcgtcctccc	gaccacccgc	ccaccccgga	gtccgcgccc	8160
ccggcttggg	tttctgcttt	gccccttccg	ccccggcccc	cctcggctcg	cggggccttt	8220
ccccccccca	cgctcgcgcc	catcccgcct	ccgcccgcgc	aaggagcggt	cgtccccggg	8280
ggcgaccgcc	gcgcggcgcc	ccgcccagaca	acagcggttc	cctcgcgcgc	acccccccgg	8340
ggcccccgcg	ccggcccacc	ccggcgcttc	acgcgccccg	cggtcgcctc	gctctcgccc	8400
togttgaact	ccctccccct	tccccgggac	ccagccgacc	acgcggccgc	ggtttccgcc	8460
gcagcggcgg	cagtgcgcgc	ttctccgggc	cttgcccccc	cgacctccgc	cgtccagacc	8520
agccccccgc	ctttggcccc	aggcccgggt	gccccgagcg	agcctttgtg	cggttgggtc	8580
gtgcccggag	gcccggttgc	ccgcgcgcgc	ccaccccagt	caccgcgcgc	caaaccgcgc	8640
gcgcggacgc	gcatccgcgc	gcgctcggtg	ccccagccgc	ccctgcccc	gcgcgccctg	8700
ccccagccgc	ccctgcccc	gcgcgccctg	ccccagccgc	ccctgcccc	gcgcgccctg	8760
ccccagccgc	ccctgcccc	gcccgcgcc	ccccagccgc	ccctgcccc	gcgcgccctg	8820
ccccagccgc	ccctgcccc	cgtaactcgg	acgtttacac	cccagagccg	agactcggtc	8880
ccgacgcccc	agtctccac	acatacaaac	accacctcc	cagtgtccgc	tgtgacttcg	8940
tgggcttctc	ctctcgcgct	tacgtcgac	tccgcgcccc	caccgcctc	gctgcttcag	9000
acgtgcaca	tctcctccga	tgatgaacac	tccgacgcag	actccctcg	tttctcggac	9060
tcggatgata	ccgaggcgct	cgacccccct	ccaccggaac	cgacttgcc	ccccgcgcgc	9120
gaaccacccg	gccattggc	cgcgcgaccac	ctccagtcac	cacactcgca	atttgcccc	9180
ttgcccgtac	aggcaaatgc	cgtcttgtcg	cgacgtacg	tacgaagcac	cggtcgtagc	9240
gcccgtggcg	ttttgatccg	cgcgtgcccgc	cgcatacaac	aacagctgca	gcggaccgcg	9300
agggcgctgt	ttcagcgacg	caacgcgcgtg	ctgaccagcc	tacatcacgt	gcgcgtgtta	9360
ctgggttag						9369

<210> 248

<211> 532

<212> PRT

<213> Herpes simplex virus

<400> 248

Met	Glu	Leu	Ser	Tyr	Ala	Thr	Thr	Leu	His	His	Arg	Asp	Val	Val	Phe
1					5				10					15	
Tyr	Val	Thr	Ala	Asp	Arg	Asn	Arg	Ala	Tyr	Phe	Val	Cys	Gly	Gly	Ser

485 490 495
 Ala Phe Asp Arg Gly Pro Ala Gly Ala Ala Gly Arg Thr Arg Thr Ala
 500 505 510
 Gly Tyr Leu Asp Ala Leu Leu Thr Val Cys Leu Ala Arg Ala Gln His
 515 520 525
 Gly Gln Ser Val
 530

<210> 249

<211> 838

<212> PRT

<213> Herpes simplex virus

<400> 249

Met Gly Pro Gly Leu Trp Val Val Met Gly Val Leu Val Gly Val Ala
 1 5 10 15
 Gly Gly His Asp Thr Tyr Trp Thr Glu Gln Ile Asp Pro Trp Phe Leu
 20 25 30
 His Gly Leu Gly Leu Ala Arg Thr Tyr Trp Arg Asp Thr Asn Thr Gly
 35 40 45
 Arg Leu Trp Leu Pro Asn Thr Pro Asp Ala Ser Asp Pro Gln Arg Gly
 50 55 60
 Arg Leu Ala Pro Pro Gly Glu Leu Asn Leu Thr Thr Ala Ser Val Pro
 65 70 75 80
 Met Leu Arg Trp Tyr Ala Glu Arg Phe Cys Phe Val Leu Val Thr Thr
 85 90 95
 Ala Glu Phe Pro Arg Asp Pro Gly Gln Leu Leu Tyr Ile Pro Lys Thr
 100 105 110
 Tyr Leu Leu Gly Arg Pro Arg Asn Ala Ser Leu Pro Glu Leu Pro Glu
 115 120 125
 Ala Gly Pro Thr Ser Arg Pro Pro Ala Glu Val Thr Gln Leu Lys Gly
 130 135 140
 Leu Ser His Asn Pro Gly Ala Ser Ala Leu Leu Arg Ser Arg Ala Trp
 145 150 155 160
 Val Thr Phe Ala Ala Pro Asp Arg Glu Gly Leu Thr Phe Pro Arg
 165 170 175
 Gly Asp Asp Gly Ala Thr Glu Arg His Pro Asp Gly Arg Arg Asn Ala
 180 185 190
 Pro Pro Pro Gly Pro Pro Ala Gly Thr Pro Arg His Pro Thr Thr Asn
 195 200 205
 Leu Ser Ile Ala His Leu His Asn Ala Ser Val Thr Trp Leu Ala Ala
 210 215 220
 Arg Gly Leu Leu Arg Thr Pro Gly Arg Tyr Val Tyr Leu Ser Pro Ser
 225 230 235 240
 Ala Ser Thr Trp Pro Val Gly Val Trp Thr Thr Gly Gly Leu Ala Phe
 245 250 255
 Gly Cys Asp Ala Ala Leu Val Arg Ala Arg Tyr Gly Lys Gly Phe Met
 260 265 270
 Gly Leu Val Ile Ser Met Arg Asp Ser Pro Pro Ala Glu Ile Ile Val
 275 280 285
 Val Pro Ala Asp Lys Thr Leu Ala Arg Val Gly Asn Pro Thr Asp Glu
 290 295 300
 Asn Ala Pro Ala Val Leu Pro Gly Pro Pro Ala Gly Pro Arg Tyr Arg
 305 310 315 320
 Val Phe Val Leu Gly Ala Pro Thr Pro Ala Asp Asn Gly Ser Ala Leu
 325 330 335
 Asp Ala Leu Arg Arg Val Ala Gly Tyr Pro Glu Glu Ser Thr Asn Tyr
 340 345 350

Ala Gln Tyr Met Ser Arg Ala Tyr Ala Glu Phe Leu Gly Glu Asp Pro
 355 360 365
 Gly Ser Gly Thr Asp Ala Arg Pro Ser Leu Phe Trp Arg Leu Ala Gly
 370 375 380
 Leu Leu Ala Ser Ser Gly Phe Ala Phe Val Asn Ala Ala His Ala His
 385 390 395 400
 Asp Ala Ile Arg Leu Ser Asp Leu Leu Gly Phe Leu Ala His Ser Arg
 405 410 415
 Val Leu Ala Gly Leu Ala Ala Arg Gly Ala Ala Gly Cys Ala Ala Asp
 420 425 430
 Ser Val Phe Leu Asn Val Ser Val Leu Asp Pro Ala Ala Arg Leu Arg
 435 440 445
 Leu Glu Ala Arg Leu Gly His Leu Val Ala Ala Ile Leu Glu Arg Glu
 450 455 460
 Gln Ser Leu Val Ala His Ala Leu Gly Tyr Gln Leu Ala Phe Val Leu
 465 470 475 480
 Asp Ser Pro Ala Ala Tyr Gly Ala Val Ala Pro Ser Ala Ala Arg Leu
 485 490 495
 Ile Asp Ala Leu Tyr Ala Glu Phe Leu Gly Gly Arg Ala Leu Thr Ala
 500 505 510
 Pro Met Val Arg Arg Ala Leu Phe Tyr Ala Thr Ala Val Leu Arg Ala
 515 520 525
 Pro Phe Leu Ala Gly Ala Pro Ser Ala Glu Gln Arg Glu Arg Ala Arg
 530 535 540
 Arg Gly Leu Leu Ile Thr Ala Leu Cys Thr Ser Asp Val Ala Ala
 545 550 555 560
 Ala Thr His Ala Asp Leu Arg Ala Ala Leu Ala Arg Thr Asp His Gln
 565 570 575
 Lys Asn Leu Phe Trp Leu Pro Asp His Phe Ser Pro Cys Ala Ala Ser
 580 585 590
 Leu Arg Phe Asp Leu Ala Glu Gly Gly Phe Ile Leu Asp Ala Leu Ala
 595 600 605
 Met Ala Thr Arg Ser Asp Ile Pro Ala Asp Val Met Ala Gln Gln Thr
 610 615 620
 Arg Gly Val Ala Ser Val Leu Thr Arg Trp Ala His Tyr Asn Ala Leu
 625 630 635 640
 Ile Arg Ala Phe Val Pro Glu Ala Thr His Gln Cys Ser Gly Pro Ser
 645 650 655
 His Asn Ala Glu Pro Arg Ile Leu Val Pro Ile Thr His Asn Ala Ser
 660 665 670
 Tyr Val Val Thr His Thr Pro Leu Pro Arg Gly Ile Gly Tyr Lys Leu
 675 680 685
 Thr Gly Val Asp Val Arg Arg Pro Leu Phe Ile Thr Tyr Leu Thr Ala
 690 695 700
 Thr Cys Glu Gly His Ala Arg Glu Ile Glu Pro Lys Arg Leu Val Arg
 705 710 715 720
 Thr Glu Asn Arg Arg Asp Leu Gly Leu Val Gly Ala Val Phe Leu Arg
 725 730 735
 Tyr Thr Pro Ala Gly Glu Val Met Ser Val Leu Leu Val Asp Thr Asp
 740 745 750
 Ala Thr Gln Gln Gln Leu Ala Gln Gly Pro Val Ala Gly Thr Pro Asn
 755 760 765
 Val Phe Ser Ser Asp Val Pro Ser Val Ala Leu Leu Leu Phe Pro Asn
 770 775 780
 Gly Thr Val Ile His Leu Leu Ala Phe Asp Thr Leu Pro Ile Ala Thr
 785 790 795 800
 Ile Ala Pro Gly Phe Leu Ala Ala Ser Ala Leu Gly Val Val Met Ile
 805 810 815

Thr Ala Ala Leu Ala Gly Ile Leu Arg Val Val Arg Thr Cys Val Pro
 820 825 830

Phe Leu Trp Arg Arg Glu
 835

<210> 250

<211> 3122

<212> PRT

<213> Herpes simplex virus

<400> 250

Met Ile Pro Ala Ala Leu Pro His Pro Thr Met Lys Arg Gln Gly Asp
 1 5 10 15
 Arg Asp Ile Val Val Thr Gly Val Arg Asn Gln Phe Ala Thr Asp Leu
 20 25 30
 Glu Pro Gly Gly Ser Val Ser Cys Met Arg Ser Ser Leu Ser Phe Leu
 35 40 45
 Ser Leu Leu Phe Asp Val Gly Pro Arg Asp Val Leu Ser Ala Glu Ala
 50 55 60
 Ile Glu Gly Cys Leu Val Glu Gly Gly Glu Trp Thr Arg Ala Ala Ala
 65 70 75 80
 Gly Ser Gly Pro Pro Arg Met Cys Ser Ile Ile Glu Leu Pro Asn Phe
 85 90 95
 Leu Glu Tyr Pro Ala Ala Arg Gly Gly Leu Arg Cys Val Phe Ser Arg
 100 105 110
 Val Tyr Gly Glu Val Gly Phe Phe Gly Glu Pro Thr Ala Gly Leu Leu
 115 120 125
 Glu Thr Gln Cys Pro Ala His Thr Phe Phe Ala Gly Pro Trp Ala Met
 130 135 140
 Arg Pro Leu Ser Tyr Thr Leu Leu Thr Ile Gly Pro Leu Gly Met Gly
 145 150 155 160
 Leu Tyr Arg Asp Gly Asp Thr Ala Tyr Leu Phe Asp Pro His Gly Leu
 165 170 175
 Pro Ala Gly Thr Pro Ala Phe Ile Ala Lys Val Arg Ala Gly Asp Val
 180 185 190
 Tyr Pro Tyr Leu Thr Tyr Tyr Ala His Asp Arg Pro Lys Val Arg Trp
 195 200 205
 Ala Gly Ala Met Val Phe Phe Val Pro Ser Gly Pro Gly Ala Val Ala
 210 215 220
 Pro Ala Asp Leu Thr Ala Ala Ala Leu His Leu Tyr Gly Ala Ser Glu
 225 230 235 240
 Thr Tyr Leu Gln Asp Glu Pro Phe Val Glu Arg Arg Val Ala Ile Thr
 245 250 255
 His Pro Leu Arg Gly Glu Ile Gly Gly Leu Gly Ala Leu Phe Val Gly
 260 265 270
 Val Val Pro Arg Gly Asp Gly Glu Gly Ser Gly Pro Val Val Pro Ala
 275 280 285
 Leu Pro Ala Pro Thr His Val Gln Thr Pro Gly Ala Asp Arg Pro Pro
 290 295 300
 Glu Ala Pro Arg Gly Ala Ser Gly Pro Pro Asp Thr Pro Gln Ala Gly
 305 310 315 320
 His Pro Asn Arg Pro Pro Asp Asp Val Trp Ala Ala Ala Leu Glu Gly
 325 330 335
 Thr Pro Pro Ala Lys Pro Ser Ala Pro Asp Ala Ala Ala Ser Gly Pro
 340 345 350
 Pro His Ala Ala Pro Pro Pro Gln Thr Pro Ala Gly Asp Ala Ala Glu
 355 360 365
 Glu Ala Glu Asp Leu Arg Val Leu Glu Val Gly Ala Val Pro Val Gly

370 375 380
 Arg His Arg Ala Arg Tyr Ser Thr Gly Leu Pro Lys Arg Arg Arg Pro
 385 390 395 400
 Thr Trp Thr Pro Pro Ser Ser Val Glu Asp Leu Thr Ser Gly Glu Arg
 405 410 415
 Pro Ala Pro Lys Ala Pro Pro Ala Lys Ala Lys Lys Lys Ser Ala Pro
 420 425 430
 Lys Lys Lys Ala Pro Val Ala Ala Glu Val Pro Ala Ser Ser Pro Thr
 435 440 445
 Pro Ile Ala Ala Thr Val Pro Pro Ala Pro Asp Thr Pro Pro Gln Ser
 450 455 460
 Gly Gln Gly Gly Gly Asp Asp Gly Pro Ala Ser Pro Ser Ser Pro Ser
 465 470 475 480
 Val Leu Glu Thr Leu Gly Ala Arg Arg Pro Pro Glu Pro Pro Gly Ala
 485 490 495
 Asp Leu Ala Gln Leu Phe Glu Val His Pro Asn Val Ala Ala Thr Ala
 500 505 510
 Val Arg Leu Ala Ala Arg Asp Ala Ala Leu Ala Arg Glu Val Ala Ala
 515 520 525
 Cys Ser Gln Leu Thr Ile Asn Ala Leu Arg Ser Pro Tyr Pro Ala His
 530 535 540
 Pro Gly Leu Leu Glu Leu Cys Val Ile Phe Phe Phe Glu Arg Val Leu
 545 550 555 560
 Ala Phe Leu Ile Glu Asn Gly Ala Arg Thr His Thr Gln Ala Gly Val
 565 570 575
 Ala Gly Pro Ala Ala Ala Leu Leu Asp Phe Thr Leu Arg Met Leu Pro
 580 585 590
 Arg Lys Thr Ala Val Gly Asp Phe Leu Ala Ser Thr Arg Met Ser Leu
 595 600 605
 Ala Asp Val Ala Ala His Arg Pro Leu Ile Gln His Val Leu Asp Glu
 610 615 620
 Asn Ser Gln Ile Gly Arg Leu Ala Leu Ala Lys Leu Val Leu Val Ala
 625 630 635 640
 Arg Asp Val Ile Arg Glu Thr Asp Ala Phe Tyr Gly Asp Leu Ala Asp
 645 650 655
 Leu Asp Leu Gln Leu Arg Ala Ala Pro Pro Ala Asn Leu Tyr Ala Arg
 660 665 670
 Leu Gly Glu Trp Leu Leu Glu Arg Ser Arg Ala His Pro Asn Thr Leu
 675 680 685
 Phe Ala Pro Ala Thr Pro Thr His Pro Glu Pro Leu Leu His Arg Ile
 690 695 700
 Gln Ala Leu Ala Gln Phe Ala Arg Gly Glu Glu Met Arg Val Glu Ala
 705 710 715 720
 Glu Ala Arg Glu Met Arg Glu Ala Leu Asp Ala Leu Ala Arg Gly Val
 725 730 735
 Asp Ser Val Ser Gln Arg Ala Gly Pro Leu Thr Val Met Pro Val Pro
 740 745 750
 Ala Ala Pro Gly Ala Gly Gly Arg Ala Pro Cys Pro Pro Ala Leu Gly
 755 760 765
 Pro Glu Ala Ile Gln Ala Arg Leu Glu Asp Val Arg Ile Gln Ala Arg
 770 775 780
 Arg Ala Ile Glu Ser Ala Val Lys Glu Tyr Phe His Arg Gly Ala Val
 785 790 795 800
 Tyr Ser Ala Lys Ala Leu Gln Ala Ser Asp Ser His Asp Cys Arg Phe
 805 810 815
 His Val Ala Ser Ala Ala Val Val Pro Met Val Gln Leu Leu Glu Ser
 820 825 830
 Leu Pro Ala Phe Asp Gln His Thr Arg Asp Val Ala Gln Arg Ala Ala

835	840	845
Leu Pro Pro Pro Pro	Leu Ala Thr Ser Pro	Gln Ala Ile Leu Leu
850	855	860
Arg Asp Leu Leu Gln Arg	Gly Gln Pro Leu Asp	Ala Pro Glu Asp Leu
865	870	875
Ala Ala Trp Leu Ser Val	Leu Thr Asp Ala Ala Thr	Gln Gly Leu Ile
885	890	895
Glu Arg Lys Pro Leu Glu	Glu Leu Ala Arg Ser Ile	His Gly Ile Asn
900	905	910
Asp Gln Gln Ala Arg Arg	Ser Ser Gly Leu Ala Glu	Leu Gln Arg Phe
915	920	925
Asp Ala Leu Asp Ala Ala	Leu Ala Gln Gln Leu Asp	Ser Asp Ala Ala
930	935	940
Phe Val Pro Ala Thr Gly	Pro Ala Pro Tyr Val Asp	Gly Gly Gly Leu
945	950	955
Ser Pro Glu Ala Thr Arg	Met Ala Glu Asp Ala Leu	Arg Gln Ala Arg
965	970	975
Ala Met Glu Ala Ala Lys	Met Thr Ala Glu Leu Ala	Pro Glu Ala Arg
980	985	990
Ser Arg Leu Arg Glu Arg	Ala His Ala Leu Glu Ala	Met Leu Asn Asp
995	1000	1005
Ala Arg Glu Arg Ala Lys	Val Ala His Asp Ala Arg	Glu Lys Phe Leu
1010	1015	1020
His Lys Leu Gln Gly Val	Leu Arg Pro Leu Pro Asp	Phe Val Gly Leu
1025	1030	1035
Lys Ala Cys Pro Ala Val	Leu Ala Thr Leu Arg Ala	Ser Leu Pro Ala
1045	1050	1055
Gly Trp Thr Asp Leu Ala	Asp Ala Val Arg Gly Pro	Pro Pro Glu Val
1060	1065	1070
Thr Ala Ala Leu Arg Ala	Asp Leu Trp Gly Leu Leu	Gly Gln Tyr Arg
1075	1080	1085
Glu Ala Leu Glu His Pro	Thr Pro Asp Thr Ala Thr	Ala Leu Ala Gly
1090	1095	1100
Leu His Pro Ala Phe Val	Val Val Leu Lys Thr Leu	Phe Ala Asp Ala
1105	1110	1115
Pro Glu Thr Pro Val Leu	Val Gln Phe Phe Ser Asp	His Ala Pro Thr
1125	1130	1135
Ile Ala Lys Ala Val Ser	Asn Ala Ile Asn Ala Gly	Ser Ala Ala Val
1140	1145	1150
Ala Thr Ala Ser Pro Ala	Ala Thr Val Asp Ala Ala	Val Arg Ala His
1155	1160	1165
Gly Ala Leu Ala Asp Ala	Val Ser Ala Leu Gly Ala	Ala Ala Arg Asp
1170	1175	1180
Pro Ala Ser Pro Leu Ser	Phe Leu Ala Val Leu Ala	Asp Ser Ala Ala
1185	1190	1195
Gly Tyr Val Lys Ala Thr	Arg Leu Ala Leu Glu Ala	Arg Gly Ala Ile
1205	1210	1215
Asp Glu Leu Thr Thr Leu	Gly Ser Ala Ala Asp Leu	Val Val Gln
1220	1225	1230
Ala Arg Arg Ala Cys Ala	Gln Pro Glu Gly Asp His	Ala Ala Leu Ile
1235	1240	1245
Asp Ala Ala Ala Arg Ala	Thr Thr Ala Ala Arg Glu	Ser Leu Ala Gly
1250	1255	1260
His Glu Ala Gly Phe Gly	Gly Leu Leu His Ala Glu	Gly Thr Ala Gly
1265	1270	1275
Asp His Ser Pro Ser Gly	Arg Ala Leu Gln Glu Leu	Gly Lys Val Ile
1285	1290	1295
Gly Ala Thr Arg Arg Arg	Ala Asp Glu Leu Glu Ala	Ala Val Ala Asp

1300 1305 1310
 Leu Thr Ala Lys Met Ala Ala Gln Arg Ala Arg Gly Ser Ser Glu Arg
 1315 1320 1325
 Trp Ala Ala Gly Val Glu Ala Ala Leu Asp Arg Val Glu Asn Arg Ala
 1330 1335 1340
 Glu Phe Asp Val Val Glu Leu Arg Arg Leu Gln Ala Leu Ala Gly Thr
 1345 1350 1355 1360
 His Gly Tyr Asn Pro Arg Asp Phe Arg Lys Arg Ala Glu Gln Ala Leu
 1365 1370 1375
 Ala Ala Asn Ala Glu Ala Val Thr Leu Ala Leu Asp Thr Ala Phe Ala
 1380 1385 1390
 Phe Asn Pro Tyr Thr Pro Glu Asn Gln Arg His Pro Met Leu Pro Pro
 1395 1400 1405
 Leu Ala Ala Ile His Arg Leu Gly Trp Ser Ala Ala Phe His Ala Ala
 1410 1415 1420
 Ala Glu Thr Tyr Ala Asp Met Phe Arg Val Asp Ala Glu Pro Leu Ala
 1425 1430 1435 1440
 Arg Leu Leu Arg Ile Ala Glu Gly Leu Leu Glu Met Ala Gln Ala Gly
 1445 1450 1455
 Asp Gly Phe Ile Asp Tyr His Glu Ala Val Gly Arg Leu Ala Asp Asp
 1460 1465 1470
 Met Thr Ser Val Pro Gly Leu Arg Arg Tyr Val Pro Phe Phe Gln His
 1475 1480 1485
 Gly Tyr Ala Asp Tyr Val Glu Leu Arg Asp Arg Leu Asp Ala Ile Arg
 1490 1495 1500
 Ala Asp Val His Arg Ala Leu Gly Gly Val Pro Leu Asp Leu Ala Ala
 1505 1510 1515 1520
 Ala Ala Glu Gln Ile Ser Ala Ala Arg Asn Asp Pro Glu Ala Thr Ala
 1525 1530 1535
 Glu Leu Val Arg Thr Gly Val Thr Leu Pro Cys Pro Ser Glu Asp Ala
 1540 1545 1550
 Leu Val Ala Cys Ala Ala Ala Leu Glu Arg Val Asp Gln Ser Pro Val
 1555 1560 1565
 Lys Asn Thr Ala Tyr Ala Glu Tyr Val Ala Phe Val Thr Arg Gln Asp
 1570 1575 1580
 Thr Ala Glu Thr Lys Asp Ala Val Val Arg Ala Lys Gln Gln Arg Ala
 1585 1590 1595 1600
 Glu Ala Thr Glu Arg Val Met Ala Gly Leu Arg Glu Ala Leu Ala Ala
 1605 1610 1615
 Arg Glu Arg Arg Ala Gln Ile Glu Ala Glu Gly Leu Ala Asn Leu Lys
 1620 1625 1630
 Thr Met Leu Lys Val Val Ala Val Pro Ala Thr Val Ala Lys Thr Leu
 1635 1640 1645
 Asp Gln Ala Arg Ser Val Ala Glu Ile Ala Asp Gln Val Glu Val Leu
 1650 1655 1660
 Leu Asp Gln Thr Glu Lys Thr Arg Glu Leu Asp Val Pro Ala Val Ile
 1665 1670 1675 1680
 Trp Leu Glu His Ala Gln Arg Thr Phe Glu Thr His Pro Leu Ser Ala
 1685 1690 1695
 Ala Arg Gly Asp Gly Pro Gly Pro Leu Ala Arg His Ala Gly Arg Leu
 1700 1705 1710
 Gly Ala Leu Phe Asp Thr Arg Arg Arg Val Asp Ala Leu Arg Arg Ser
 1715 1720 1725
 Leu Glu Glu Ala Glu Ala Glu Trp Asp Glu Val Trp Gly Arg Phe Gly
 1730 1735 1740
 Arg Val Arg Gly Gly Ala Trp Lys Ser Pro Glu Gly Phe Arg Ala Met
 1745 1750 1755 1760
 His Glu Gln Leu Arg Ala Leu Gln Asp Thr Thr Asn Thr Val Ser Gly

1765 1770 1775
 Leu Arg Ala Gln Pro Ala Tyr Glu Arg Leu Ser Ala Arg Tyr Gln Gly
 1780 1785 1790
 Val Leu Gly Ala Lys Gly Ala Glu Arg Ala Glu Ala Val Glu Glu Leu
 1795 1800 1805
 Gly Ala Arg Val Thr Lys His Thr Ala Leu Cys Ala Arg Leu Arg Asp
 1810 1815 1820
 Glu Val Val Arg Arg Val Pro Trp Glu Met Asn Phe Asp Ala Leu Gly
 1825 1830 1835 1840
 Gly Leu Leu Ala Glu Phe Asp Ala Ala Ala Asp Leu Ala Pro Trp
 1845 1850 1855
 Ala Val Glu Glu Phe Arg Gly Ala Arg Glu Leu Ile Gln Tyr Arg Met
 1860 1865 1870
 Gly Leu Tyr Ser Ala Tyr Ala Arg Ala Gly Gly Gln Thr Gly Ala Gly
 1875 1880 1885
 Ala Glu Ser Ala Pro Ala Pro Leu Leu Val Asp Leu Arg Ala Leu Asp
 1890 1895 1900
 Ala Arg Ala Arg Ala Ser Ser Ser Pro Glu Gly His Glu Val Asp Pro
 1905 1910 1915 1920
 Gln Leu Leu Arg Arg Arg Gly Glu Ala Tyr Leu Arg Ala Gly Gly Asp
 1925 1930 1935
 Pro Gly Pro Leu Val Leu Arg Glu Ala Val Ser Ala Leu Asp Leu Pro
 1940 1945 1950
 Phe Ala Thr Ser Phe Leu Ala Pro Asp Gly Thr Pro Leu Gln Tyr Ala
 1955 1960 1965
 Leu Cys Phe Pro Ala Val Thr Asp Lys Leu Gly Ala Leu Leu Met Arg
 1970 1975 1980
 Pro Glu Ala Ala Cys Val Arg Pro Pro Leu Pro Thr Asp Val Leu Glu
 1985 1990 1995 2000
 Ser Ala Pro Thr Val Thr Ala Met Tyr Val Leu Thr Val Val Asn Arg
 2005 2010 2015
 Leu Gln Leu Ala Leu Ser Asp Ala Gln Ala Ala Asn Phe Gln Leu Phe
 2020 2025 2030
 Gly Arg Phe Val Arg His Arg Gln Ala Thr Trp Gly Ala Ser Met Asp
 2035 2040 2045
 Ala Ala Ala Glu Leu Tyr Val Ala Leu Val Ala Thr Thr Leu Thr Arg
 2050 2055 2060
 Glu Phe Gly Cys Arg Trp Ala Gln Leu Gly Trp Ala Ser Gly Ala Ala
 2065 2070 2075 2080
 Ala Pro Arg Pro Pro Pro Gly Pro Arg Gly Ser Gln Arg His Cys Val
 2085 2090 2095
 Ala Phe Asn Glu Asn Asp Val Leu Val Ala Leu Val Ala Gly Val Pro
 2100 2105 2110
 Glu His Ile Tyr Asn Phe Trp Arg Leu Asp Leu Val Arg Gln His Glu
 2115 2120 2125
 Tyr Met His Leu Thr Leu Glu Arg Ala Phe Glu Asp Ala Ala Glu Ser
 2130 2135 2140
 Met Leu Phe Val Gln Arg Leu Thr Pro His Pro Asp Ala Arg Ile Arg
 2145 2150 2155 2160
 Val Leu Pro Thr Phe Leu Asp Gly Gly Pro Pro Thr Arg Gly Leu Leu
 2165 2170 2175
 Phe Gly Thr Arg Leu Ala Asp Trp Arg Arg Gly Lys Leu Ser Glu Thr
 2180 2185 2190
 Asp Pro Leu Ala Pro Trp Arg Ser Ala Leu Glu Leu Gly Thr Gln Arg
 2195 2200 2205
 Arg Asp Val Pro Ala Leu Gly Lys Leu Ser Pro Ala Gln Ala Leu Ala
 2210 2215 2220
 Ala Val Ser Val Leu Gly Arg Met Cys Leu Pro Ser Ala Ala Leu Ala

2225 2230 2235 2240
 Ala Leu Trp Thr Cys Met Phe Pro Asp Asp Tyr Thr Glu Tyr Asp Ser
 2245 2250 2255
 Phe Asp Ala Leu Leu Ala Ala Arg Leu Glu Ser Gly Gln Thr Leu Gly
 2260 2265 2270
 Pro Ala Gly Gly Arg Glu Ala Ser Leu Pro Glu Ala Pro His Ala Leu
 2275 2280 2285
 Tyr Arg Pro Thr Gly Gln His Val Ala Val Leu Ala Ala Ala Thr His
 2290 2295 2300
 Arg Thr Pro Ala Ala Arg Val Thr Ala Met Asp Leu Val Leu Ala Ala
 2305 2310 2315 2320
 Val Leu Leu Gly Ala Pro Val Val Val Ala Leu Arg Asn Thr Thr Ala
 2325 2330 2335
 Phe Ser Arg Glu Ser Glu Leu Glu Leu Cys Leu Thr Leu Phe Asp Ser
 2340 2345 2350
 Arg Pro Gly Gly Pro Asp Ala Ala Leu Arg Asp Val Val Ser Ser Asp
 2355 2360 2365
 Ile Glu Thr Trp Ala Val Gly Leu Leu His Thr Asp Leu Asn Pro Ile
 2370 2375 2380
 Glu Asn Ala Cys Leu Ala Ala Gln Leu Pro Arg Leu Ser Ala Leu Ile
 2385 2390 2395 2400
 Ala Glu Arg Pro Leu Ala Asp Gly Pro Pro Cys Leu Val Leu Val Asp
 2405 2410 2415
 Ile Ser Met Thr Pro Val Ala Val Leu Trp Glu Ala Pro Glu Pro Pro
 2420 2425 2430
 Gly Pro Pro Asp Val Arg Phe Val Gly Ser Glu Ala Thr Glu Glu Leu
 2435 2440 2445
 Pro Phe Val Ala Thr Ala Gly Asp Val Leu Ala Ala Ser Ala Ala Asp
 2450 2455 2460
 Ala Asp Pro Phe Phe Ala Arg Ala Ile Leu Gly Arg Pro Phe Asp Ala
 2465 2470 2475 2480
 Ser Leu Leu Thr Gly Glu Leu Phe Pro Gly His Pro Val Tyr Gln Arg
 2485 2490 2495
 Pro Leu Ala Asp Glu Ala Gly Pro Ser Ala Pro Thr Ala Ala Arg Asp
 2500 2505 2510
 Pro Arg Asp Leu Ala Gly Gly Asp Gly Gly Ser Gly Pro Glu Asp Pro
 2515 2520 2525
 Ala Ala Pro Pro Ala Arg Gln Ala Asp Pro Gly Val Leu Ala Pro Thr
 2530 2535 2540
 Leu Leu Thr Asp Ala Thr Thr Gly Glu Pro Val Pro Pro Arg Met Trp
 2545 2550 2555 2560
 Ala Trp Ile His Gly Leu Glu Glu Leu Ala Ser Asp Asp Ala Gly Gly
 2565 2570 2575
 Pro Thr Pro Asn Pro Ala Pro Ala Leu Leu Pro Pro Pro Ala Thr Asp
 2580 2585 2590
 Gln Ser Val Pro Thr Ser Gln Tyr Ala Pro Arg Pro Ile Gly Pro Ala
 2595 2600 2605
 Ala Thr Ala Arg Glu Thr Arg Pro Ser Val Pro Pro Gln Gln Asn Thr
 2610 2615 2620
 Gly Arg Val Pro Val Ala Pro Arg Asp Asp Pro Arg Pro Ser Pro Pro
 2625 2630 2635 2640
 Thr Pro Ser Pro Pro Ala Asp Ala Ala Leu Pro Pro Pro Ala Phe Ser
 2645 2650 2655
 Gly Ser Ala Ala Ala Phe Ser Ala Ala Val Pro Arg Val Arg Arg Ser
 2660 2665 2670
 Arg Arg Thr Arg Ala Lys Ser Arg Ala Pro Arg Ala Ser Ala Pro Pro
 2675 2680 2685
 Glu Gly Trp Arg Pro Pro Ala Leu Pro Ala Pro Val Ala Pro Val Ala

2690 2695 2700
 Ala Ser Ala Arg Pro Pro Asp Gln Pro Pro Thr Pro Glu Ser Ala Pro
 2705 2710 2715 2720
 Pro Ala Trp Val Ser Ala Leu Pro Leu Pro Pro Gly Pro Ala Ser Ala
 2725 2730 2735
 Arg Gly Ala Phe Pro Ala Pro Thr Leu Ala Pro Ile Pro Pro Pro Pro
 2740 2745 2750
 Ala Glu Gly Ala Val Val Pro Gly Gly Asp Arg Arg Arg Gly Arg Arg
 2755 2760 2765
 Gln Thr Thr Ala Gly Pro Ser Pro Thr Pro Pro Arg Gly Pro Ala Ala
 2770 2775 2780
 Gly Pro Pro Arg Arg Leu Thr Arg Pro Ala Val Ala Ser Leu Ser Ala
 2785 2790 2795 2800
 Ser Leu Asn Ser Leu Pro Ser Pro Arg Asp Pro Ala Asp His Ala Ala
 2805 2810 2815
 Ala Val Ser Ala Ala Ala Ala Val Pro Pro Ser Pro Gly Leu Ala
 2820 2825 2830
 Pro Pro Thr Ser Ala Val Gln Thr Ser Pro Pro Pro Leu Ala Pro Gly
 2835 2840 2845
 Pro Val Ala Pro Ser Glu Pro Leu Cys Gly Trp Val Val Pro Gly Gly
 2850 2855 2860
 Pro Val Ala Arg Arg Pro Pro Pro Gln Ser Pro Ala Thr Lys Pro Ala
 2865 2870 2875 2880
 Ala Arg Thr Arg Ile Arg Ala Arg Ser Val Pro Gln Pro Pro Leu Pro
 2885 2890 2895
 Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln
 2900 2905 2910
 Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro
 2915 2920 2925
 Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro Leu Pro Gln Pro Pro
 2930 2935 2940
 Leu Pro Pro Val Thr Arg Thr Leu Thr Pro Gln Ser Arg Asp Ser Val
 2945 2950 2955 2960
 Pro Thr Pro Glu Ser Pro Thr His Thr Asn Thr His Leu Pro Val Ser
 2965 2970 2975
 Ala Val Thr Ser Trp Ala Ser Ser Leu Ala Leu His Val Asp Ser Ala
 2980 2985 2990
 Pro Pro Pro Ala Ser Leu Leu Gln Thr Leu His Ile Ser Ser Asp Asp
 2995 3000 3005
 Glu His Ser Asp Ala Asp Ser Leu Arg Phe Ser Asp Ser Asp Asp Thr
 3010 3015 3020
 Glu Ala Leu Asp Pro Leu Pro Pro Glu Pro His Leu Pro Pro Ala Asp
 3025 3030 3035 3040
 Glu Pro Pro Gly Pro Leu Ala Ala Asp His Leu Gln Ser Pro His Ser
 3045 3050 3055
 Gln Phe Gly Pro Leu Pro Val Gln Ala Asn Ala Val Leu Ser Arg Arg
 3060 3065 3070
 Tyr Val Arg Ser Thr Gly Arg Ser Ala Leu Ala Val Leu Ile Arg Ala
 3075 3080 3085
 Cys Arg Arg Ile Gln Gln Gln Leu Gln Arg Thr Arg Arg Ala Leu Phe
 3090 3095 3100
 Gln Arg Ser Asn Ala Val Leu Thr Ser Leu His His Val Arg Met Leu
 3105 3110 3115 3120
 Leu Gly

<210> 251

<211> 625

<212> DNA

<213> Herpes simplex virus

<400> 251

cggacgctcc	ccccggctcg	ccagccccc	cacccccga	acatcgcggc	ggacccgagg	60
agtttgaggg	cgccggggac	ggcgaacccc	ccgatgacga	cgacagcgcc	accggcctcg	120
ccttcggaac	tccgaacccc	aacaaaccac	cccccgcgcg	ccccggggcc	atccgcccc	180
cgctcccgcc	aggaattctt	gggcccgtcg	cccccaacac	gcctcgcccc	cccgcccaag	240
ctcccgctaa	ggacatgccc	tcgggcccc	caccccaaca	catccccctg	ttctgggtcc	300
taacggcctc	ccctgctcta	gatatcctct	ttatcatcag	caccaccatc	cacacggcgg	360
cgttcgtttg	tctggtcgcc	ttggcagcac	aactttggcg	cggccgggcg	gggcgcaggc	420
gatacgcgca	cccagcgctg	cgttacgtat	gtctgccacc	cgagcgggat	tagggggtgg	480
ggtgggggcg	agaaacgatg	aaggacggga	aagggaacag	cgaccaaatg	ccacgataag	540
aacaataaac	ctgtgacgtc	aatcggatat	gtgagtttgg	ttgtgttttg	tgggactgct	600
cgtgcgcgtc	gtgcgcgtcg	tgccg				625

<210> 252

<211> 2100

<212> DNA

<213> Herpes simplex virus

<400> 252

atgcacgcca	tcgctcccag	gttgettcct	ctttttgttc	tttctggtct	tccggggaca	60
cgcggggggt	cggtgtcccc	cggaccaatt	aateccccca	acagcgatgt	tgttttcccg	120
ggaggttccc	ccgtggctca	atattgttat	gcctatcccc	ggttggaaga	tcccgggccc	180
ttgggttcgg	cggacgccc	gcggcaagac	ctgccccggc	gcgtcgcccg	tcacgagccc	240
ctgggcccgt	cgttcctcac	ggggggggctg	gttttgcctg	cgccgcccgt	acgcggattt	300
ggcgacacca	acgcaacgta	tgccggcccgt	gtgacgtact	accggctcac	ccgcgcctgc	360
cgtcagccca	tcctccttcg	gcagtatgga	gggtgtcgcg	gcggcgagcc	gcccgcacca	420
aagacgtgcg	ggtcgtacac	gtacacgtac	cagggcgggc	ggcctccgac	ccggtacgct	480
ctcgtaaagt	cttccctgct	ggtgccgatc	tgggaccgcg	ccgcggagac	attcgagtac	540
cagatcgaac	tcggcgggca	gctgcacgtg	ggtctgttgt	gggtagaggt	gggcggggag	600
ggccccggcc	ccaccgcccc	cccacaggcg	gcgcgtgcgg	agggcggccc	gtgcgtcccc	660
ccggtccccg	cgggcccggc	gtggcgctcg	gtgcccccg	tatggtatcc	cgcgcccaac	720
cccgggtttc	gtggcctgcg	tttccgggag	cgctgtctgc	ccccacagac	gcccgcgcgc	780
cccagcgacc	taccacgcgt	cgtttttgct	ccccagagcc	tgctggtggg	gattacgggc	840
cgcacgttta	ttcggatggc	acgaaccaag	gaagacgtcg	gggtcctgcc	gccccattgg	900
gccccggggg	ccctagatga	cggctccgtac	gcccccttcc	caccccgcgc	gcggtttcga	960
cgcgccttgc	ggacagaccc	cgagggggtc	gaccccgacg	ttcggggccc	ccgaaccggg	1020
cggcgccctc	tgcccttgac	cgaggacacg	tcctccgatt	cgctacgctc	cgctccggag	1080
aagacgcccc	tcctctgtgc	ggccaccgcc	atggcaccct	cagtcgaccc	aagcgcgga	1140
ccgaccgccc	ccgcaaccac	tactcccccc	gacgagatgg	ccacacaagc	cgcaaccggtc	1200
gccgttacgc	cggaggaaac	ggcagtcgcc	tccccgccc	cgactgcac	cgtggagtcg	1260
tcgccactcc	ccgccgcggc	ggcggaacg	cccggggccg	ggcacacgaa	caccagcagc	1320
gcctccgcag	cgaaaacgcc	ccccaccaca	ccagccccca	cgaccccccc	gcccacgtct	1380
accacgcga	ccccccgccc	cacgaactcg	gggccccaaa	caacccctcc	cggaccgcga	1440
accccgggtc	cgttgggcgc	ctccgcgcgc	cccacggccg	attccccctc	caccgcctcg	1500
cccccgcta	ccgcgcgggg	gcccctggcc	gccaacgttt	cggtcgcgcg	gaccaccgcc	1560
acgcccggaa	cccggggcac	cgcccgtacc	cccccaacg	acccaaagac	gcaccacac	1620
ggaccgcggy	acgttcccc	cggtctgcca	gccccccac	ccccgaaca	tcgcggcgga	1680
cccaggaggt	ttgaggcgcg	cggggacggc	gaaccccccg	aggacgacga	cagcgccacc	1740
ggcctcgctc	tccgaactcc	gaacccccac	aaaccacccc	ccgcgcgccc	cgggcccatc	1800
cgcccacgc	tcccgcagg	aattcttggg	ccgctcgccc	ccaacacgcc	tcgccccccc	1860
gcccaagctc	ccgctaagga	catgccctcg	ggccccacac	cccaacacat	ccccctgttc	1920
tggttcctaa	cggcctcccc	tgctctagat	atcctcttta	tcacagcac	caccatccac	1980
acggcgcggt	tcgtttgtct	ggtcgccctg	gcagcacaac	tttggcgcg	ccgggcgggg	2040
cgcaggcgat	acgcgcaccc	gagcgtgcgt	tacgtatgtc	tgccaccgca	gcgggattag	2100

<210> 253
 <211> 156
 <212> PRT
 <213> Herpes simplex virus

<400> 253
 Asp Ala Pro Pro Gly Ser Pro Ala Pro Pro Pro Pro Glu His Arg Gly
 1 5 10 15
 Gly Pro Glu Glu Phe Glu Gly Ala Gly Asp Gly Glu Pro Pro Glu Asp
 20 25 30
 Asp Asp Ser Ala Thr Gly Leu Ala Phe Arg Thr Pro Asn Pro Asn Lys
 35 40 45
 Pro Pro Pro Ala Arg Pro Gly Pro Ile Arg Pro Thr Leu Pro Pro Gly
 50 55 60
 Ile Leu Gly Pro Leu Ala Pro Asn Thr Pro Arg Pro Pro Ala Gln Ala
 65 70 75 80
 Pro Ala Lys Asp Met Pro Ser Gly Pro Thr Pro Gln His Ile Pro Leu
 85 90 95
 Phe Trp Phe Leu Thr Ala Ser Pro Ala Leu Asp Ile Leu Phe Ile Ile
 100 105 110
 Ser Thr Thr Ile His Thr Ala Ala Phe Val Cys Leu Val Ala Leu Ala
 115 120 125
 Ala Gln Leu Trp Arg Gly Arg Ala Gly Arg Arg Arg Tyr Ala His Pro
 130 135 140
 Ser Val Arg Tyr Val Cys Leu Pro Pro Glu Arg Asp
 145 150 155

<210> 254
 <211> 699
 <212> PRT
 <213> Herpes simplex virus

<400> 254
 Met His Ala Ile Ala Pro Arg Leu Leu Leu Leu Phe Val Leu Ser Gly
 1 5 10 15
 Leu Pro Gly Thr Arg Gly Gly Ser Gly Val Pro Gly Pro Ile Asn Pro
 20 25 30
 Pro Asn Ser Asp Val Val Phe Pro Gly Gly Ser Pro Val Ala Gln Tyr
 35 40 45
 Cys Tyr Ala Tyr Pro Arg Leu Asp Asp Pro Gly Pro Leu Gly Ser Ala
 50 55 60
 Asp Ala Gly Arg Gln Asp Leu Pro Arg Arg Val Val Arg His Glu Pro
 65 70 75 80
 Leu Gly Arg Ser Phe Leu Thr Gly Gly Leu Val Leu Leu Ala Pro Pro
 85 90 95
 Val Arg Gly Phe Gly Ala Pro Asn Ala Thr Tyr Ala Ala Arg Val Thr
 100 105 110
 Tyr Tyr Arg Leu Thr Arg Ala Cys Arg Gln Pro Ile Leu Leu Arg Gln
 115 120 125
 Tyr Gly Gly Cys Arg Gly Gly Glu Pro Pro Ser Pro Lys Thr Cys Gly
 130 135 140
 Ser Tyr Thr Tyr Thr Tyr Gln Gly Gly Gly Pro Pro Thr Arg Tyr Ala
 145 150 155 160
 Leu Val Asn Ala Ser Leu Leu Val Pro Ile Trp Asp Arg Ala Ala Glu
 165 170 175
 Thr Phe Glu Tyr Gln Ile Glu Leu Gly Gly Glu Leu His Val Gly Leu
 180 185 190
 Leu Trp Val Glu Val Gly Gly Glu Gly Pro Gly Pro Thr Ala Pro Pro

660 665 670
Gln Leu Trp Arg Gly Arg Ala Gly Arg Arg Arg Tyr Ala His Pro Ser
675 680 685
Val Arg Tyr Val Cys Leu Pro Pro Glu Arg Asp
690 695

<210> 255
<211> 15
<212> PRT
<213> HSV-2

<400> 255
Ile Trp Thr Gly Asn Pro Arg Thr Ala Pro Arg Ser Leu Ser Leu
5 10 15

<210> 256
<211> 15
<212> PRT
<213> HSV-2

<400> 256
Pro Arg Thr Ala Pro Arg Ser Leu Ser Leu Gly Gly His Thr Val
5 10 15

<210> 257
<211> 9
<212> PRT
<213> HSV-2

<400> 257
Arg Thr Ala Pro Arg Ser Leu Ser Leu
5

<210> 258
<211> 15
<212> PRT
<213> HSV-2

<400> 258
Pro Ala Pro Pro Ala Val Pro Val Asp Ala His Arg Ala Pro Arg
5 10 15

<210> 259
<211> 15
<212> PRT
<213> HSV-2

<400> 259
Ser Gly Arg Ala Ala Arg Pro Arg Ala Ala Val Ala Pro Arg Val
5 10 15

<210> 260
<211> 15
<212> PRT

<213> HSV-2

<400> 260

Ala Gln Val Ser Ser Gly Pro Gly Gly Gly Gly Leu Pro Gln Ser
5 10 15

<210> 261

<211> 15

<212> PRT

<213> HSV-2

<400> 261

Tyr Ala Gly Arg Met Thr Tyr Ile Ala Thr Gly Ala Leu Leu Ala
5 10 15

<210> 262

<211> 15

<212> PRT

<213> HSV-2

<400> 262

Met Thr Tyr Ile Ala Thr Gly Ala Leu Leu Ala Arg Phe Asn Pro
5 10 15

<210> 263

<211> 9

<212> PRT

<213> HSV

<400> 263

Thr Tyr Ile Ala Thr Gly Ala Leu Leu
5

<210> 264

<211> 15

<212> PRT

<213> HSV-2

<400> 264

Ala Arg Leu His Pro His Ser Ala His Pro Ala Phe Ala Asp Val
5 10 15

<210> 265

<211> 9

<212> PRT

<213> HSV

<400> 265

His Pro His Ser Ala His Pro Ala Phe
5

<210> 266

<211> 15

<212> PRT

<213> HSV

<400> 266

Ala Ser Thr Arg Ser His Ala Pro Ser Ala His Arg Ala Asp Ser
5 10 15

<210> 267

<211> 15

<212> PRT

<213> HSV

<400> 267

Met Ser Val Arg Gly His Ala Val Arg Arg Arg Arg Ala Ser Thr
5 10 15